



中国科学院大学

University of Chinese Academy of Sciences

硕士学位论文

适用于大数据时代的星系面亮度拟合程序

作者姓名: 陈宓

指导教师: 沈世银 研究员 中国科学院上海天文台

Rafael S de Souza senior lecturer University of Hertfordshire

学位类别: 理学硕士

学科专业: 天体物理

培养单位: 中国科学院上海天文台

2024年6月

Galaxy Profile Fitting Software under Big Data Era

**A thesis submitted to
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Master of Science
in Astrophysics**

By

Chen Mi

Supervisor: Professor Shen Shiyin

Professor Rafael S. de Souza

Shanghai Astronomical Observatory, Chinese Academy of Sciences

June, 2024

中国科学院大学 学位论文原创性声明

本人郑重声明：所提交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学 学位论文授权使用声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘要

星系是一个由气体、尘埃、恒星和暗晕组成的复杂生态系统，显示出包括大小、形态在内的各异物理性质。理解星系形态和结构的多样性是研究星系演化的一个基本目标，而星系面亮度拟合正是对星系形态与结构进行量化的重要方法之一。然而，在 LSST、CSST、Euclid 等大型巡天项目已经或即将开始工作的时代，传统的面亮度拟合方法已经无法适应海量数据的产生。为了解决传统方法速度慢、需要人工干预的特点，利用深度学习建立新映射、GPU 加速面亮度拟合过程等方法应运而生。

为此，我们编写了 GALMOSS，一个基于 Python 与 PyTorch，使用 GPU 加速进行大样本二维星系面亮度轮廓拟合的程序。通过将星系数据集加载为张量进行多维矩阵计算来利用 GPU 并行化，GALMOSS 能够满足未来星系大样本巡天的高计算需求，使其能够适应全新的巡天时代。GALMOSS 内置了七种被广泛使用的星系面亮度轮廓模型，例如 Sérsic、指数盘、Ferrer、King、高斯和 Moffat，并允许用户进行任意数量模型共同拟合。在性能测评中，我们使用单张 NVIDIA A100 GPU 对来自 Sloan 数字巡天 (SDSS) g 波段的 8,289 个星系进行了测试，在大约 10 分钟内完成了 Sérsic 轮廓的单成分拟合。在与 GALFIT 结果高度一致的同时，我们发现 GALMOSS 的拟合结果比 GALFIT 快出六倍（GALFIT 使用 2.2GHz Intel Xeon Silver 4210）。

关键词：星系: 综合-方法: 数据分析-方法: 统计-GPU 计算

Abstract

Galaxies are complex ecosystems comprising diverse components such as gas, dust, stellar, and dark matter halos, display a wide variety of physical properties including their sizes and shapes. Understanding the diversity of galaxy morphologies and structures is a fundamental goal in studying galaxy evolution, and galaxy profile fitting is one of the important methods to quantify them. It is well known that With the expanding volume of data from astronomical surveys such as the Legacy Survey of Space and Time (LSST), Chinese Survey Space Telescope (CSST) and Euclid, significant challenges emerge in galaxy profile fitting tasks. Traditional approaches face difficulties in processing the vast number of galaxies efficiently and automatically. To address these challenges, deep learning strategies and parallel gradient-based methodology are increasingly being employed.

As one of the solutions, we introduce GALMOSS, a Python-based, torch-powered tool for two-dimensional fitting of galaxy profiles. By seamlessly enabling GPU parallelization, GALMOSS meets the high computational demands of large-scale galaxy surveys, placing galaxy profile fitting in the LSST-era. It incorporates widely used profiles such as the Sérsic, Exponential disk, Ferrer, King, Gaussian, and Moffat profiles, and allows for the easy integration of more complex models. Tested on 8,289 galaxies from the Sloan Digital Sky Survey (SDSS) g-band with a single NVIDIA A100 GPU, GALMOSS completed classical Sérsic profile fitting in about 10 minutes. Benchmark tests show that GALMOSS achieves computational speeds that are $6 \times$ faster than GALFIT (with CPU –2.2GHz Intel Xeon Silver 4210).

Keywords: galaxies: general – methods: data analysis – methods: statistical – GPU computing

目 录

第 1 章 绪论	1
1.1 轮廓拟合	3
1.1.1 各类面亮度轮廓模型简介	4
1.1.2 轮廓拟合的历史发展	6
1.1.3 轮廓拟合的应用	8
1.2 大数据时代与深度学习	10
1.2.1 巡天项目	10
1.2.2 深度学习以及天文领域的应用	12
1.3 利用不同方法的面亮度轮廓拟合程序	13
1.3.1 传统方法	13
1.3.2 深度学习方法	14
1.3.3 GPU 加速方法	15
1.4 本文概览	16
1.4.1 本文研究目标	16
1.4.2 本文结构	17
第 2 章 GALMOSS: 基于 GPU 加速的星系面亮度拟合程序	19
2.1 工作流程	19
2.2 图像的生成	20
2.2.1 从参数到模型图像	20
2.2.2 PSF 卷积	26
2.2.3 GPU: 图像生成过程的加速原理	27
2.3 图像的拟合	29
2.3.1 拟合优度	29
2.3.2 参数优化方法	29
2.3.3 GPU: 参数优化的加速原理	32
2.4 参数的误差估计	33
2.4.1 协方差矩阵	34
2.4.2 自举重采样	34

第 3 章 性能评估：面亮度轮廓重拟合实验	37
3.1 数据	37
3.1.1 SDSS 的测光图像	37
3.1.2 MANGA 的测光与形态星表	38
3.1.3 数据集的制作	39
3.2 实验设置	40
3.2.1 GALMOSS 重拟合实验设置	40
3.2.2 统计指标	41
3.3 拟合结果与讨论	42
3.3.1 拟合准确性与讨论	43
3.3.2 拟合误差的评估	48
3.3.3 拟合速度的评估	49
第 4 章 总结、讨论与未来工作展望	51
4.1 总结	51
4.2 讨论与未来展望	52
附录 A galmoSS 使用示例	55
A.1 单轮廓使用示例	55
A.2 双轮廓使用示例	57
A.3 多轮廓建模	58
A.4 自定义轮廓示例	62
参考文献	65
致谢	75
作者简历及攻读学位期间发表的学术论文与研究成果	77

图形列表

1.1 哈勃序列。图源 https://sci.esa.int 。	2
1.2 一个 SB0 星系，NGC 4665 在 r 波段的面亮度轮廓图像，取自 Erwin 等 (2007)。	3
1.3 图左所显示的是 N3348 的面亮度轮廓以及其与 Core Sérsic 轮廓拟合的结果。虚线为 Sérsic 轮廓的面亮度轮廓，可以看到 N3348 在中心有不 _{符合} Sérsic 轮廓的扁平区域。图右是一个符合经典 Sérsic 面亮度轮廓的星系 N5831。	5
1.4 J162123.19+322056.4 星系面亮度拟合的结果。左边为数据、模型与残差。右边为结果的一维投影。由于二维图像无法很好地进行模型、数据和残差的同时展示，在评估结果时通常会对拟合结果进行投影，在原有拟合结果的基础上增加一维轮廓进行展示。	7
1.5 椭圆星系的 B 波段中央面亮度 $\mu_{0,B}$ 与 (a) B 波段星等 M_B 和与 (b) Sérsic 指数 n 的线性关系。c 图为星系星等与 Sérsic 指数 n 的关系。其中，黑点与三角形为源自不同文章的 dE 星系、大星形为 Coma dE 星系，小星形为中等至明亮的 E 星系，空心圆为“power-law” E 星系，实心圆为含核区流量的星系。三图的拟合结果均表现为图中的黑线，分别为 (a) $M_B = 0.67\mu_{0,B} - 29.5$ ，(b) $\mu_0 = 22.8 - 14.1 \log(n)$ ，(c) $M_B = -9.4 \log(n) - 14.3$ 。具体的阐述与数据来源详见 Oswald 等 (2013)。	8
1.6 类似的，椭圆星系的 B 波段星等 M_B 与 (a) 平均有效面亮度 $\langle \mu \rangle_e$ 和与 (b) 有效面亮度 μ_e 和与 (c) 有效半径 R_e 的关系。其中，曲线关系由图 1.5 所展示的关系导出。具体的阐述与数据来源详见 Oswald 等 (2013)。	8
1.7 现代望远镜每晚能释放的数据量级。图源 Kremer 等 (2017)。	11
2.1 GALMOSS 的工作流程：从参数初始化和图像处理到拟合过程和数据产出的导出。最终的数据产出包括拟合模型和残差，默认全部以 FITS 格式存储。	20
2.2 Sérsic 轮廓随指数因子 n 的变化。图源 Shen 等 (2022)。	21
2.3 盒状星系与盘状星系。	22
2.4 修正的 Ferrer 随斜率参数的变化，图源 Peng (2003)。	23
2.5 King 轮廓随核半径、幂律因子的变化，图源 Peng (2003)。	24
2.6 Moffat 轮廓随指数因子的变化，图源 Peng (2003)。	25
2.7 GPU 矩阵计算加速的原理，图源 https://docs.nvidia.com/ 。	28

2.8 GPU 矩阵计算对算力的使用, 图源 https://docs.nvidia.com/ 。 ·····	28
2.9 反向传播的动态梯度图, 图源 https://pytorch.org/ 。 ·····	32
2.10 在自举重采样法中对星系图像进行重采样。 ·····	34
3.1 星系图像数据集的实例。 ·····	40
3.2 测试星系的 loss 曲线。 ·····	41
3.3 GALMOSS 拟合结果的视觉检查。 ·····	43
3.4 对 8,329 个星系进行拟合后的 χ_v 直方图。 ·····	44
3.5 此图比较了所选 $\sim 8,000$ 个 SDSS 星系的测量 Sérsic 剖面参数, 以及其与 GALFIT 结果 (MPP-VAC 星表) 比较的 R^2 决定系数。子图分别显示了星等 (m)、方位角 (θ)、轴比 (q)、有效半径 (r_e) 和 Sérsic 指数 (n) 的拟合结果。点的颜色说明了它们分布的密度。蓝线显示 galfit 结果完美符合 galmooss 结果的参考线。顶部和右侧的直方图分别说明了 galmooss 和 galfit 测量值的分布。此图显示, GALMOSS 拟合的所有参数都与 GALFIT 拟合的结果具有高度的一致性。 ·····	45
3.6 S0 星系、椭圆星系不同来源拟合值的直方图。紫色代表了 S0 星系 Sersic 指数 n 的分布, 红色代表了椭圆星系 Sersic 指数 n 的分布; 而虚线则为 GALFIT 拟合的结果, 实线为 GALMOSS 拟合的结果。由图可得, 椭圆星系对拟合 n 值的不一致贡献最大。 ·····	46
3.7 对椭圆星系进行不同迭代数拟合的结果。随着迭代数的增加, GALMOSS 的结果逐渐向 GALFIT 逼近。 ·····	47
3.8 使用样本中随机选择的 10 个星系对 GALMOSS 内置的两种拟合不确定性估计方法进行比较: 协方差矩阵 (用蓝色误差棒表示) 和自举法 (用红色误差条表示)。通常, 使用自举重采样法得出的不确定性大于使用协方差矩阵获得的不确定性。 ·····	48
3.9 GALMOSS 和 GALFIT 的运行时间比较, 以 GALFIT 运行时间为标准进行归一化。结果表明, 当批数量稍大于 10 的时候, GALMOSS 开始超越 GALFIT 的速度。当批数量为 1000 的时候, GALMOSS 将 GALFIT 的速度提升了大约六倍。 ·····	50
4.1 Mask-RCNN 的流程图。Mask-RCNN 在提取特征图之后, 通过 Align layer 将特征图映射为相同大小。随后, 该网络分为两个分支, 第一个分支用于对此特征图进行分类, 第二个分支用于产生掩膜图像。 ···	52
4.2 Mask-RCNN 的训练对比图。第一列为原图 (为了凸显出中央星系所在位置, 图像经过拉伸); 第二列为 SExtractor 所产生的掩膜图像; 第三列为 Mask-RCNN 产生的掩膜图像, 第四列为利用 Mask-RCNN 的掩膜图像去除干扰天体后的星系流量图像。 ·····	53
4.3 ·····	54

A.1 J162123.19+322056.4 星系 g 波段面亮度拟合的结果。左边为数据、模型与残差。右边为结果的一维投影。 ·····	56
A.2 J100247.00+042559.8 星系 g 波段面亮度拟合的结果。左边为数据、模型与残差。右边为结果的一维投影。 ·····	57
A.3 未经过 PSF 卷积的模型图像，包含中央星系、背景星系与两个点源。	61
A.4 经过 PSF 卷积的模型图像，包含中央星系、背景星系与两个点源。 ·	62

表格列表

1.1 不同形态类型的 K 波段旋涡星系所展现的形态参数。数据来源 (Graham 等, 2008)	9
3.1 星系分类方法	39
3.2 GALMOSS 拟合结果的评估	44

第 1 章 绪论

星系作为宇宙的重要组成部分,是由气体、尘埃、恒星和暗晕相互影响组成的复杂系统。气体和尘埃在引力的作用下坍缩形成恒星 (Jeans, 1901; Larson, 1969),而作为恒星形成与演化的产物,超大质量黑洞及其吸积盘会通过辐射或动力学驱动反馈来改变气体与尘埃的温度和化学成分 (Mineshige 等, 1997; Toyouchi 等, 2020),这一过程又将反过来促进或抑制该区域未来的恒星形成 (Kim 等, 2011; Martín-Navarro 等, 2018)。与其他组分不同,暗物质目前无法被直接观测,但其与其他可见组分相互作用的过程可以通过引力来追溯。比如在星系演化的早期,暗物质较为集中的区域是新恒星的“种子”(Springel 等, 2003; Dayal 等, 2018),并进一步影响星系与星团的形成 (Governato 等, 1998)。在这个复杂的生态系统中,各种物理过程以不同星系组分为载体影响星系形成与演化的路径,从而使星系在演化的各个阶段呈现出各异的性质,并导向不同的演化结果。

星系形态和结构是星系的基本性质之一,包括星系尺度、形态以及各类子结构,如核球、盘、棒、团块、旋臂等。作为星系最直观的物理性质,星系形态和结构与其他性质的关系是研究星系形成与演化的重要一环。比如,星系的大小和星系光度、星系颜色、恒星质量和红移等性质显著相关 (Shen 等, 2003; Ferguson 等, 2004; Williams 等, 2010; Mosleh 等, 2020)。星系旋臂的数量、强弱与恒星形成过程有紧密联系 (Hart 等, 2017; Yu 等, 2021)。星系形态与颜色和所在小、大尺度环境的关系有助于理解星系所处环境对自身演化过程的影响 (Foex 等, 2009; Park 等, 2009; Liu 等, 2019)。

对星系形态与结构的观测有着悠久的历史,多种不同的方法被开发与使用。星系的光谱能量分布拟合可以表征不同恒星星族在光谱中的占比。这些恒星星族在星系中拥有对应的区域,各异的占比也使得星系呈现出不同的结构 (Walcher 等, 2011; Pacifici 等, 2023; Conroy, 2013)。除此之外,星系的数值模拟则通过建立三维模型,模拟从气体到大质量星系的完整动力学过程和演化历史,实现了对星系的形态与结构理论角度上的分解和研究 (Schwarzschild, 1979; Zhu 等, 2018)。与其他方法相比,天文图像则是历史最悠久,且目前仍被广泛运用的方法。星系的形态以流量的形式被图像展示,并由此延伸出了对星系的目视形态分类。同时,通过研究图像中星系的光度分布,延伸出了非参数的星系形态分析方法与参数的星系面亮度轮廓拟合。

目视形态分类使用肉眼分辨星系形态的方式对星系进行归类。由于分类结果与星系的形态表现直接相关,因此目视形态分类的可解释性较强,也非常直观。在诸多形态分类的星表中, Hubble (1926) 首次发表的哈勃序列是目前运用最广泛的一种。图1.1展示了当下时间、40 亿以及 110 亿年前,位于哈勃序列中的星系的形态特征。音叉状的哈勃序列将星系分为旋涡星系 (S)、椭圆星系 (E) 以及透镜星系 (S0)。音叉图中的分类还可以被重新归类为早、晚型两种类型。

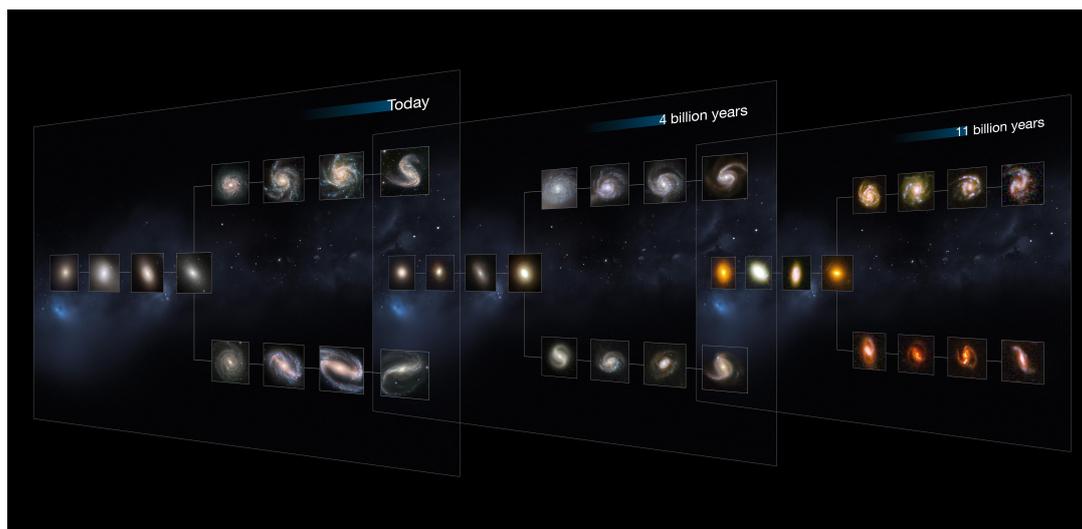


图 1.1 哈勃序列。图源<https://sci.esa.int>。

椭圆星系与透镜星系常被归类为早型星系，而旋涡星系、不规则星系则常被归类为晚型星系。虽然这很容易与演化的顺序联系起来，但实际上目前的观测的事实却是相反的结果，宇宙的早期更多地由旋涡和不规则星系组成 (Abraham 等, 2001)。而椭圆星系则是由晚型星系并合而成 (George, 2017)。定性的目视分类直观便捷，但却缺乏定量的数学描述。分立的类别在描述变化连续的星系形态时，经常会遇见定义模糊不清，或分类困难的情况。因此，各种定量描述星系形态的方法也应运而生。

非参数形态分析法利用特殊设计的数学表达式对星系图像进行计算，并以此得到相应的数值参数。这些数值参数可以评估星系某种特征的强弱，并以此为基础进一步进行星系形态的分类或者统计分析。经典的非参数指标包括使用表示包含星系 20% 与 80% 处的半径长度进行定义的聚集度 $C = 5 \log_{10}(r_{80}/r_{20})$ ；使用星系原图像 (I) 与经过 180° 旋转后的图像 (R) 定义的不对称度 $A = \text{abs}(I - R)/I$ 、使用星系原图像 (I) 与经过平滑的图像 (B) 定义的团块度 $S = (I - B)/I$ 等 (Bershady 等, 2000; Schade 等, 1995; Conselice, 2000)。除了 CAS 参数之外，基尼系数 (G)、矩指数 (M20) 也被广泛用于表示星系的总体结构与并合星系的认证 (Lotz 等, 2004; Zamojski 等, 2007; Lotz 等, 2008; Thompson 等, 2015; Cibinel 等, 2015; Drew 等, 2020)。此方法通过统一的公式消除了主观性带来的不稳定性。而且，由于不需要先行地选择某种模型，此方法也不会产生基于假设的系统偏差，得到的结果比较稳健。但数值参数作为一种评价指标，它可以通过星系图像计算得到，但不能在仅拥有参数数值的情况下进行对应星系图像的映射。因此，非参数形态分析法并不直观。

为了兼顾量化性与直观性，轮廓拟合方法应运而生。

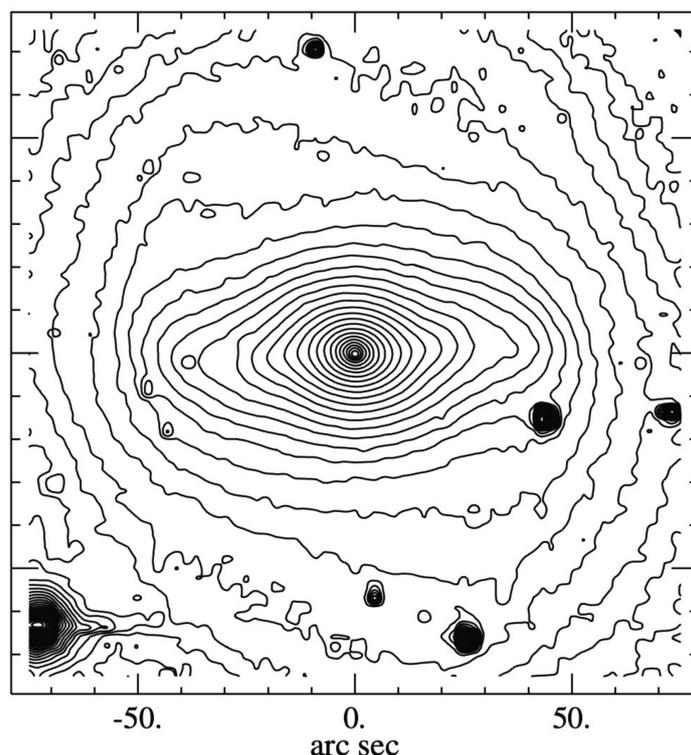


图 1.2 一个 SB0 星系，NGC 4665 在 r 波段的面亮度轮廓图像，取自 Erwin 等 (2007)。

1.1 轮廓拟合

星系的面亮度轮廓是我们对星系的结构和形态进行研究的重要方法。星系在某波段所辐射的光在投影在天球上时，其分布可以用辐射强度 $I(x, y)$ 来描述，单位为 $(\text{erg s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1})$ 。星系与星团的面亮度轮廓 $I(x, y)$ 通常随着半径 r 的增大而减小，这是由恒星的密度分布所决定的。在引力束缚的作用下，恒星的分布密度通常由内而外逐渐减小。因此，大部分由恒星贡献的亮度，其分布共享着相同的负相关关系。但由于投影效应的作用，含有一定倾角的星系，其在不同方位角的半径将呈现出一定的变化，并表现为星系不同的方位角和轴比（如图1.2）。因此在这些星系中，面亮度轮廓将与椭圆的等效半径有相关性。

面亮度轮廓的拟合即使用一些已有定论的面亮度轮廓模型去拟合所观测的星系图像。在拟合过程中，不仅可以得到精准的椭圆等照度线（即确认了星系中心、轴比与方位角），还能得知面亮度轮廓与椭圆等效半径具体的量化关系。如图1.2所示，在靠近星系中心的位置，面亮度变化的速度较快，等照度轮廓线的距离非常近；在远离星系中心的位置，面亮度变化明显放缓，等照度轮廓线的距离也相较星系中心变大。而通过使用具体的模型来拟合这些关系，所建立的模型图像可以直观再现星系的流量分布，所得到的模型参数还能进行进一步的统计研究。

1.1.1 各类面亮度轮廓模型简介

早在“星云”、星系的世纪之争结束之前 (Hubble, 1929), Plummer (1911) 与 Reynolds (1913) 就以他们关于“星云”中恒星径向分布的轮廓模型开启了天文领域关于面亮度轮廓的研究。随后, 一系列研究工作以椭圆星系为起始点开始展开。

de Vaucouleurs (1948) 用法语发表了对明亮“星云”——椭圆“星云” NGC 3115 (E7), NGC 3379 (E0), NGC 4649 (E2) 以及螺旋“星云”: NGC 4594 (Sa) 的面亮度研究, 并提出了 **de Vaucouleurs 轮廓**, 即 $r^{1/4}$ 律, 开启了它在亮椭圆星系面亮度拟合领域半世纪的统治。但在当时, 矮星系的面亮度轮廓作为另一条赛道, **King 轮廓**比 de Vaucouleurs 轮廓更为适合 (Hodge, 1961b,a)。之后, 虽然 King 模型 (King, 1962, 1966) 是为描述星团的面亮度轮廓而构建, 却常被使用于拟合矮星系。但 King 模型在此领域的流行很快被**指数轮廓**所打破, 因为 King 模型被证实其潮汐截断的设置与平坦的内轮廓并不适用于所有的椭圆星系 (Lauer, 1985), 且指数模型被发现同样适合描述矮星系 (De Vaucouleurs, 1959; Hodge, 1971; Faber 等, 1983; Binggeli 等, 1984)。虽 King 模型在矮星系领域折戟, 但它并未就此淘汰。King 模型在球状星团、矮球状星系、星系的卫星星系等拥有平坦面亮度轮廓核心的天体中依然得到广泛的使用 (Tripathi 等, 2023; Burkert, 2015; Majewski 等, 2002)。就此, 明亮、暗弱两种椭圆星系两条赛道都拥有了自己适合的面亮度轮廓模型。然而, de Vaucouleurs 轮廓的统治地位却使两条赛道的统一经历了漫长的过程。由于它所适合的范围非常广泛, 即使 de Vaucouleurs 轮廓是一个经验模型, 但在 20 世纪的天文研究中却被称为“定律”。而与此同时, **Sérsic 轮廓**即 $r^{1/n}$ 律的首次提出并非是为了椭圆星系, 而是为了对盘星系的面亮度轮廓进行建模。Sérsic (1963) 认为, 盘星系是由一个具有不同于 $r^{1/4}$ 面亮度聚集方式的盘以及符合 $r^{1/4}$ 律的核球所组成。因此, Sérsic 轮廓的提出并未在一开始就对 de Vaucouleurs 轮廓产生冲击。尽管早在 1966 年 King 就在其关于 King 轮廓文章的讨论中提出, de Vaucouleurs 轮廓似乎仅描述了某种特定的面亮度聚集的形式 (King, 1966), 经过大半个世纪, 历经 (Capaccioli 等, 1985; Capaccioli, 1987, 1989; Caon 等, 1993; D’Onofrio 等, 1994) 等工作的努力, 人们才开始广泛意识到, 椭圆星系具有一系列不同的面亮度聚集形式, 可以用 Sérsic 轮廓中的指数因子进行量化 (Trujillo 等, 2001)。

时至今日, Sérsic 轮廓作为 de Vaucouleurs 轮廓更广义的定义 ($r^{1/4}$ 律推广到 $r^{1/n}$ 律), 被广泛应用于椭圆星系的面亮度轮廓建模与盘星系的成分分解等。Oswalt 等 (2013) 在其综述中指出, 面亮度轮廓与其对应的 Sérsic 模型的偏差总是由一些额外存在的子成分组成, 比如在椭圆星系的中心出现的核星团或大质量椭圆星系中心存在的扁平的核区流量、盘星系的截断等。在盘星系中, 环、棒、旋臂等丰富的子结构也都超出了 Sérsic 轮廓的表达能力。而随着天文图像质量的提高, 针对这些结构的面亮度模型被陆续开发出来, 适用于各类拥有迥异面亮度分布的天体。

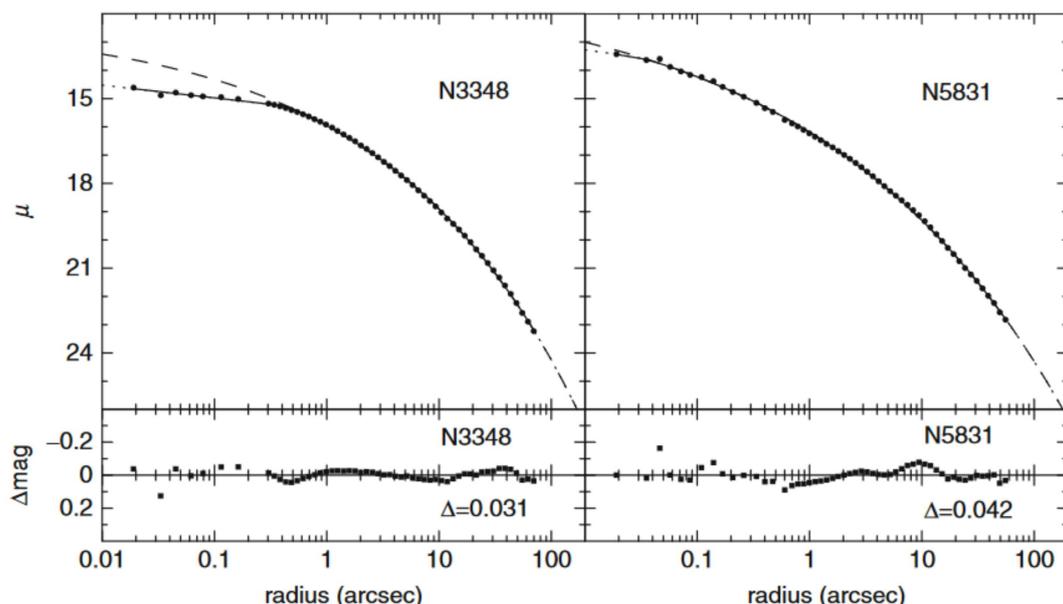


图 1.3 图左所显示的是 N3348 的面亮度轮廓以及其与 Core Sérsic 轮廓拟合的结果。虚线为 Sérsic 轮廓的面亮度轮廓，可以看到 N3348 在中心有不符合 Sérsic 轮廓的扁平区域。图右是一个符合经典 Sérsic 面亮度轮廓的星系 N5831。

Core-Sérsic 轮廓是目前针对亮椭圆星系的最佳模型，它们通常包含有 Sérsic 模型所无法很好拟合的平坦核区（如图1.3所示）。然而，Core-Sérsic 模型的产生也并非一蹴而就。在 Reynolds 模型、de Vaucouleurs 模型与 King 模型先后被证实无法对这些特殊的亮椭圆星系进行准确描述后 (King 等, 1972; Lauer, 1985)，哈勃空间望远镜的高质量图像给天文学家们提供了新的可能。Grillmair 等 (1994); Kormendy 等 (1994); Lauer 等 (1995) 等人提出了一种双幂律的面亮度轮廓公式，称为 Nuker law，非常好地再现了去除其他流量后的核区面亮度分布。然而如上所述，Nuker Law 并非服务于全半径的面亮度分布，仅仅符合小范围内的核区流量。同时，Nuker Law 也完全没有考虑与大半径处所符合的轮廓模型的衔接（如当时流行的 $r^{1/4}$ 律）。种种原因导致了 Nuker Law 在推广时的崩溃 (Graham 等, 2003; Trujillo 等, 2004; Dullo 等, 2012)。在 2003 年，Graham 提出了 Core-Sérsic 模型 (Graham 等, 2003)。该模型对 Sérsic 模型进行了修改，使其在小半径时符合幂律，以出现相对平坦的核区流量分布。与核区成分相类似，**截断的指数盘轮廓** (Erwin 等, 2007) 同样是观测技术进步的产物。人们发现，用单一的指数盘已经无法准确地描述盘状星系 (Pohlen 等, 2006; Erwin 等, 2007)。除了一些中间不存在截断的盘星系可以使用单指数盘之外，还存在一些截断前后斜率不相同的星系，需要一个双指数模型来描述。在使用截断的指数盘模型进行拟合后，根据截断的指数盘过渡斜率的大小，可以将盘状星系划分为没有断裂的单指数盘、截断后斜率更陡的盘，以及截断后斜率较缓的盘 (Laine 等, 2014)。

当观测分辨率允许时，盘星系丰富的子结构也可以被建模。如使用在大半径处急剧下降的高斯函数来近似描述棒的面亮度分布的 **Ferrer 轮廓**、描述团块面

亮度分布的**高斯轮廓**，模拟星系环、尘埃带的**截断函数**，以及模拟旋臂 (Peng 等, 2002) 的**傅里叶模式**等。除此之外，核棒与核盘也同样可以被建模 (de Zeeuw 等, 1991; Rest 等, 2001)。

除了针对天体的轮廓之外，还有一些特殊的面亮度轮廓模型，如对点扩散函数进行建模的**Moffat 轮廓** (Moffat, 1969)，对天空背景进行模拟的**天空轮廓**等。这些特殊的轮廓模型也大多会被面亮度轮廓拟合程序所收纳。

在 GALMOSS 中，我们囊括了包括 Sérsic 轮廓、指数盘轮廓、Moffat 轮廓、King 轮廓、Ferrer 轮廓、高斯轮廓、天空轮廓在内的面亮度轮廓模型。详细的公式会在 2.2.1 章进行进一步的阐述。

1.1.2 轮廓拟合的历史发展

星系面亮度拟合经历了由一维到二维的发展。在此方法发展的初期，人们通常将观测得到的星系图像沿着长轴或短轴进行平均，从而得到与方位角无关的一维星系面亮度轮廓，并总结其轮廓分布的规律。一系列轮廓模型公式在此时被提出并沿用至今，如最早被应用于早型星系的 de Vaucouleurs 轮廓 (de Vaucouleurs, 1948)，描述盘星系的指数轮廓 (De Vaucouleurs, 1959) 等。Sérsic (1963) 将以上轮廓集成为 $r^{1/n}$ 律 (Sérsic 轮廓)，并成为如今最为应用的轮廓模型之一。除了使用单一轮廓模型去描述整个星系的面亮度轮廓分布，Kormendy (1977); Burstein (1979); Kent (1985); Send (1982) 等人还基于一维星系轮廓进行了星系结构分解。

随着面亮度轮廓拟合领域的发展，一维的轮廓拟合自身存在的许多局限性已经无法满足对拟合精度的要求 (Peng 等, 2002, 2010)。比如，在将二维的星系头像投影为一维轮廓的时候，选择长轴、短轴或椭圆投影将会对投影结果产生系统误差。而且当星系的等照度线较为复杂时，不可能完全在避免轮廓模型参与的前提下做到正确的投影。其次，一维轮廓拟合因为难以考虑点扩散函数卷积的问题，无法将望远镜分辨率纳入考量范围，会引入较大的系统误差。由于投影的过程是将二维信息压缩为一维信息的过程，其总是伴随着信息量的损失。流量平均的过程如何考虑误差的平均，非对称结构如何影响一维轮廓的分布等等，这些都是值得考量的问题。比如，棒结构的存在将会在一维轮廓上体现出类核球的表现，还容易引起混淆。因此当进行包含不对称结构的星系结构分解时，一维轮廓拟合无法进行准确的结构分解。再次，投影的过程要求各个子结构共享轮廓中心。然而在实际的星系图像中，它们可能并不如此 (Lange 等, 2015)。

二维的轮廓拟合方法可以较好地避免一维拟合方法所出现的如上问题。二维面亮度轮廓的拟合进行星系图像与轮廓模型之间像素对像素的比较，并支持使用卷积点扩散函数来模拟视宁度对观测的影响。由于支持像素对像素拟合，其还支持星系不同的子结构使用不同的中心、方位角、轴比等几何特征，支持加入非对称的子结构。这些细节都能够提升拟合的精确度。1995 年，Byun 首次将二维面亮度轮廓拟合用于模拟星系，并发现二维拟合的结果远远优于一维拟合的结果 (Byun 等, 1995)。由此开始，二维轮廓拟合逐渐取代一维轮廓拟合，成为

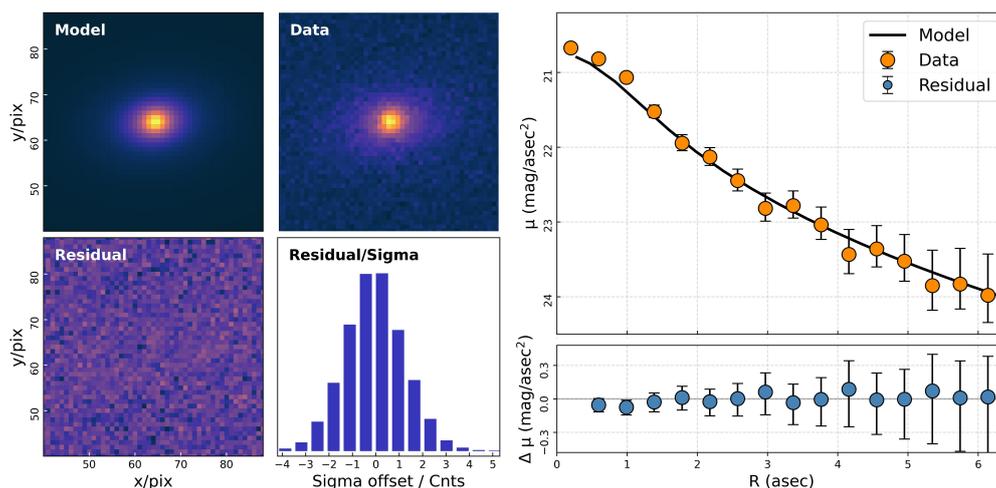


图 1.4 J162123.19+322056.4 星系面亮度拟合的结果。左边为数据、模型与残差。右边为结果的一维投影。由于二维图像无法很好地进行模型、数据和残差的同时展示，在评估结果时通常会对拟合结果进行投影，在原有拟合结果的基础上增加一维轮廓进行展示。

了目前的主流方法 (de Jong, 1996; Van Dokkum 等, 2008; Paulino-Afonso 等, 2016; George 等, 2024)。

但目前,关于一维和二维图像分析的相对优点仍然存在一些争论。如 Savorgnan 等 (2016) 在其利用有关星系分解进行黑洞质量与宿主星系的研究中发现, 由于对同一星系所使用的轮廓模型具有显著差异, 不同研究的结果并没有达成一致的结论。在独立进行了一维和二维面亮度轮廓分解后, Savorgnan 得出结论: 至少在对大型、紧邻的星系进行轮廓拟合时, 一维分解比二维分解更具优势。最后, 他们释放了可以考虑系统误差的一维面亮度轮廓分解程序, 并认为比目前流行的仅考虑统计误差的二维程序更稳健。除此之外, 二维的面亮度轮廓在进行等照度积分时需要进行计算代价较大的数值估计 (Robotham 等, 2017)。并且, 复杂的轮廓形状 (如带有旋转的等照度轮廓) 非常难以模拟。虽然原则上这些可以被合并到二维的面亮度轮廓拟合中 (Peng 等, 2010), 但由于操作复杂, 在真正的实践中很少被采用。

因此在实际应用中, 视研究需求的不同, 可以在一维与二维的轮廓拟合中选取更为适合的种类。比如, 在具有高分辨率的星系图像中 (如上文提到的大型、紧邻的星系), 点扩散效应并不十分重要, 且复杂的几何效应非常显著。在这种情况下, 使用一维的轮廓拟合将更加合适。而在大样本的星系轮廓拟合需求中则恰恰相反。由于二维拟合的方式不需要将数据进行预先的投影, 因此使用更加方便, 便于进行自动化处理。而且相较于一维方法, 二维拟合可以在更少的人工参与下进行拟合。

图 1.4 展示了一个对 J162123.19+322056.4 星系进行二维面亮度轮廓拟合的例子。左边展示了较为直接的图像结果, 分别为星系原图、模型图像、残差图像以及残差的分布。而右边则是投影后的一维轮廓展示。由图可知, 轮廓拟合除了得到的轮廓参数外, 通过最佳轮廓参数产生的模型图像和残差图像以及其一维

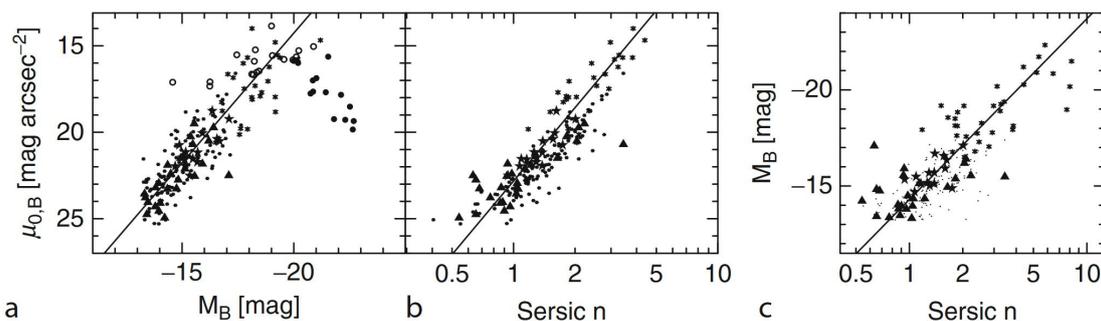


图 1.5 椭圆星系的 B 波段中央面亮度 $\mu_{0,B}$ 与 (a) B 波段星等 M_B 和与 (b) Sérsic 指数 n 的线性关系。c 图为星系星等与 Sérsic 指数 n 的关系。其中，黑点与三角形为源自不同文章的 dE 星系、大星形为 Coma dE 星系，小星形为中等至明亮的 E 星系，空心圆为“power-law” E 星系，实心圆为含核区流量的星系。三图的拟合结果均表现为图中的黑线，分别为 (a) $M_B = 0.67\mu_{0,B} - 29.5$ ，(b) $\mu_0 = 22.8 - 14.1 \log(n)$ ，(c) $M_B = -9.4 \log(n) - 14.3$ 。具体的阐述与数据来源详见 [Oswalt 等 \(2013\)](#)。

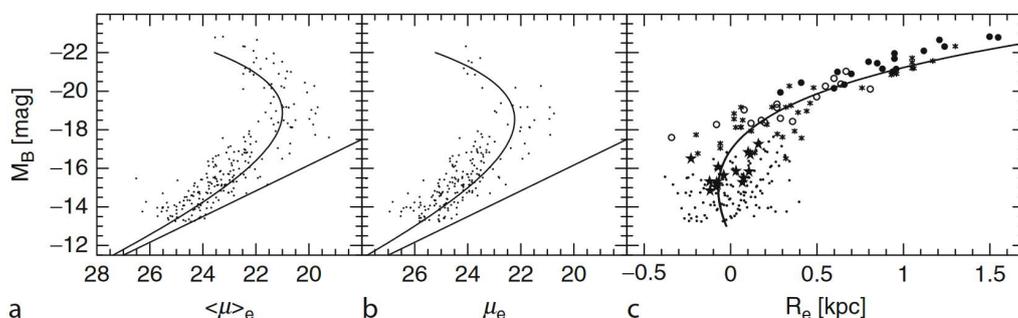


图 1.6 类似的，椭圆星系的 B 波段星等 M_B 与 (a) 平均有效面亮度 $\langle \mu \rangle_e$ 和与 (b) 有效面亮度 μ_e 和与 (c) 有效半径 R_e 的关系。其中，曲线关系由图 1.5 所展示的关系导出。具体的阐述与数据来源详见 [Oswalt 等 \(2013\)](#)。

投影还能直观地表达出星系的形态特征与拟合优度。因此，星系的面亮度轮廓结合了可解释性和量化性的双重优点。

1.1.3 轮廓拟合的应用

轮廓拟合利用模型描述了星系的面亮度轮廓，量化了星系的尺度以及面亮度的分布特征，并被大量运用于星系结构与形态的研究。椭圆星系的三类双参数关系就是其中一种。

第一类是表达了矮椭圆星系与椭圆星系统一性质的三种线性关系 (如图 1.5)。Caldwell (1983) 与 Bothun 等 (1986) 发现，在面亮度低于 $M_B \approx -20.5$ 时，星系的光度与中央面亮度之间有连续的线性关系。这一关系分别在 Binggeli 等 (1984) 与 Binggeli 等 (1991) 使用 King 模型、Jerjen 等 (1998) 与 Graham 等 (2003) 使用 Sérsic 轮廓的工作中得到进一步的证实。类似的，Ichikawa 等 (1986) 提出星系的光度与光度的聚集度之间也存在相似的线性关系 (聚集度通常使用 Sérsic 轮廓的指数 n 来量化)。而通过这两个线性关系，还可以简单地推导出星系光度聚集度与中央面亮度之间的线性关系。

由上述线性关系出发，第二类涉及有效半径 R_e 、有效面亮度 μ_e 或有效半径内的平均面亮度 $\langle \mu \rangle_e$ ，都为曲线关系（如图1.6）。Binggeli 等 (1984) 与 Capaccioli 等 (1991) 最先由观测数据导出星系光度与有效面亮度的曲线关系，并由 Graham 等 (2003) 对此关系做出解释。与之相类似的，还有同样由 Binggeli 等 (1984) 阐述的尺度-光度关系，且被外推到透镜星系与旋涡星系的盘成分上 (de Jong, 1996; Simard 等, 2011)。不同形态的星系，其大小-亮度关系的数值也不尽相同 (Shen 等, 2003; Bernardi 等, 2014)。然而在过去，这样的非线性关系导致了长久以来将矮、巨椭圆星系分为两种完全不同的星系。人们在选择样本时，往往倾向于筛去中间的星系，使得两分类的特征变得更加明显 (Oswalt 等, 2013)。Graham 在他们 2006 年的工作中指出，大质量椭圆星系样本将比普通椭圆星系样本具有斜率更大的尺寸-光度关系，而普通椭圆星系样本又将比矮椭圆星系样本具有斜率更大的尺度-光度关系，因此尺度-光度关系为曲线关系 (Graham 等, 2006)。但当选择效应出现时，人们往往会得到线性关系的结果（如 Kormendy (1977) 认为，亮椭圆星系的有效面亮度 $SB\mu_e$ 与有效半径 R_e 之间存在线性关系）。在图1.5.a 的大亮度端，我们还能观察到非常明显的偏离，即截断关系的一种。这通常被认为是由 $M_B = -20.5$ 附近 Sérsic 轮廓与 core-Sérsic 的转变导致。一般认为，在这些拥有大光度的 core-Sérsic 星系中，通常有一个由于并合事件产生的核，而且它是部分耗散的。

Type	Sérsic n	R_e/h	$\log(B/D)$
Sa	$2.56^{+2.79}_{-0.79}$	$0.31^{+0.20}_{-0.20}$	$-0.34^{+0.40}_{-0.40}$
Sab	$2.45^{+1.27}_{-0.75}$	$0.24^{+0.22}_{-0.22}$	$-0.54^{+0.53}_{-0.53}$
Sb	$2.00^{+1.62}_{-0.76}$	$0.21^{+0.15}_{-0.07}$	$-0.60^{+0.28}_{-0.49}$
Sbc	$1.87^{+1.64}_{-0.75}$	$0.21^{+0.11}_{-0.09}$	$-0.82^{+0.28}_{-0.42}$
Sc	$1.78^{+2.18}_{-0.79}$	$0.22^{+0.27}_{-0.09}$	$-1.06^{+0.43}_{-0.34}$
Scd	$1.18^{+0.89}_{-0.49}$	$0.19^{+0.10}_{-0.06}$	$-1.23^{+0.75}_{-0.28}$
Sd	$1.80^{+0.49}_{-1.22}$	$0.24^{+0.07}_{-0.10}$	$-1.06^{+0.16}_{-0.50}$
Sdm	$0.79^{+0.24}_{-0.13}$	$0.19^{+0.14}_{-0.07}$	$-1.49^{+0.36}_{-0.36}$
Sm	$0.40^{+0.09}_{-0.09}$	$0.23^{+0.01}_{-0.01}$	$-1.57^{+0.01}_{-0.01}$

表 1.1 不同形态类型的 K 波段旋涡星系所展现的形态参数。数据来源 (Graham 等, 2008)

星系面亮度拟合还被大量运用于盘星系的结构分解。de Vaucouleurs (1958) 最早尝试将星系的面亮度轮廓分解为核球成分与盘成分。他分别从长轴与短轴对星系面亮度轮廓进行投影，并发现 M31 的核球成分的面亮度分布符合 de Vaucouleurs 轮廓。Freeman (1970) 将旋涡星系与 S0 星系分为核球成分与盘成分，并首次阐述了盘成分的指数性质，即指数轮廓。随后，人们逐渐发现盘星系的核球成分并不总是能用 $R^{1/4}$ 律来描述 (Shaw 等, 1989; Wainscoat 等, 1989)。于是，使用 Sérsic 轮廓与指数轮廓对盘星系进行结构分解成为了较为常用的做法。如表1.1，为Graham 等 (2008) 发表的关于 K 波段旋涡星系结构参数结果。类似的，

一系列大样本工作围绕核球、盘分解的结果展开。如Allen 等 (2006) 证实星系的颜色双峰分布是由核球、盘成分而非不同的星族而引起。Kormendy 等 (2004) 发现 Sérsic 轮廓的指数因子可用于区分核球与伪核球。Kennedy 等 (2016) 通过对核球、盘颜色以及 Sérsic 轮廓的指数因子的关系研究解释了 Sérsic 指数对波长的依赖。而望远镜精度的提升，星系棒、团块、旋臂、环等子结构也加入了星系的结构分解 (Weinzirl 等, 2009; Gao 等, 2017; Cava 等, 2018)。

1.2 大数据时代与深度学习

目前，基于地面和空间的天文观测的正在不断建设和发展。从伽马射线、X射线、紫外线、光学、红外线到射电波段的全波段观测正在将天文学带入前所未有的大数据时代。Zhang 等 (2015) 提出，大数据时代有 4V 特征：Volume (数据量)、Variety (数据的复发性)、Velocity (数据出产的速度) 和 Value (数据的高质量)。

与其他学科类似，数据“海啸”对传统天文研究方法产生了较大的冲击，以前小精尖的数据处理方式已经无法跟上数据产生的速度，必须与深度学习等技术进行交叉融合。下面，我们先列举一些著名巡天项目，然后介绍深度学习及其在天文中的应用。

1.2.1 巡天项目

巡天项目通常不存在特定的观测目标，而是拥有一个更宽泛的科学目标，且这个科学目标往往需要以大样本的统计为研究基础。因此，依照科学目标的不同，不同的巡天项目会拥有对应的波段与对应的天区。与高精度的特定目标观测不同，巡天项目往往具有很大的视场，在每次观测都会产生大量的数据。如图1.7所示，现代望远镜每晚能释放的数据量级已经由 GB 量级快速增长到了 TB 量级。而在总数据量级上，已经由 2MASS、SDSS、LAMOST、Gaia 的 TB 量级到 PanSTARRS、FAST、Euclid、CSST、LSST 的 PB 量级涨到 SKA 的 EB 量级。

SDSS

Sloan Digital Sky Survey (SDSS)，是目前影响力最大的巡天项目之一。它首先通过 u、g、r、i、z 五个波段的成像观测进行选源，再使用多目标光谱仪进行批量的光谱观测。自 2000 年始，SDSS 先后经历了四个具有不同的科学目标的观测阶段，释放了超过 600TB 的惊人数据量，完成了包括 Sloan 遗珍巡天、Sloan 超新星巡天、APO 银河演化实验、重子振荡光谱巡天等等计划。其中，2014 年开始观测的 MANGA 项目对 10,000 个经过筛选的近邻星系样本进行了积分视场光谱仪的观测，得到三维测光数据 (CUBE)。CUBE 数据可以更深入地分析其结构，例如径向速度和恒星形成区域等。2020 年，SDSS 正式进入第五阶段的观测，拥有首个在整个天区提供多历元光学和红外光谱的设施。

DESI 与 DESI Legacy Imaging Surveys

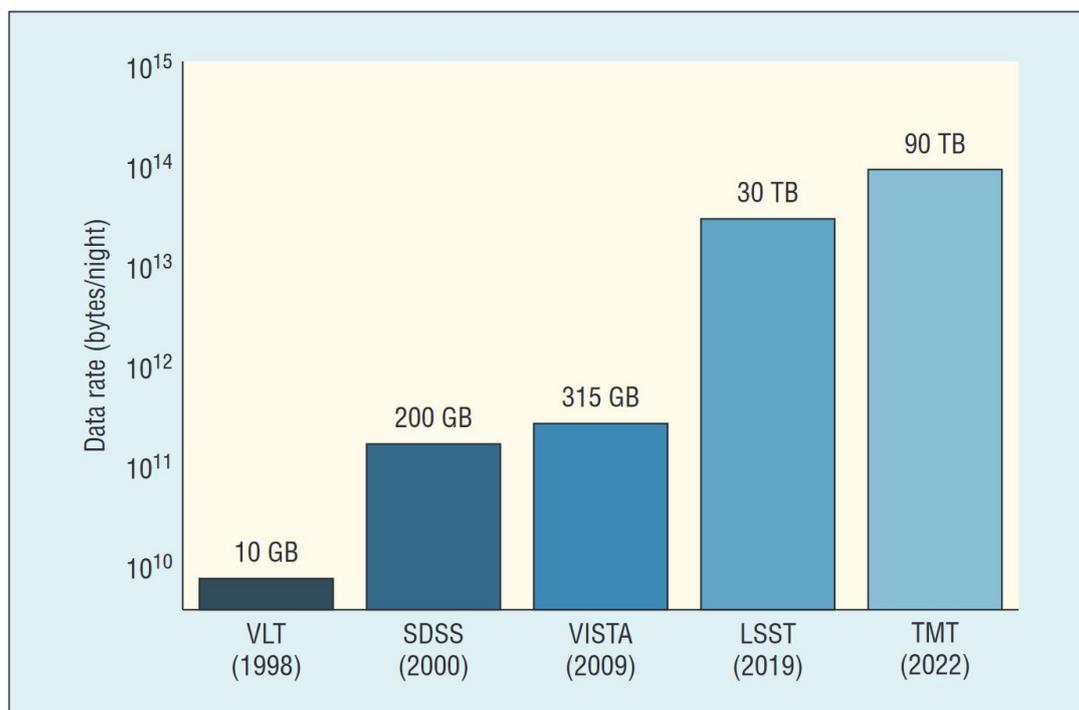


图 1.7 现代望远镜每晚能释放的数据量级。图源 Kremer 等 (2017)。

Dark Energy Spectroscopic Instrument (DESI), 即暗能量光谱仪, 是一个五年的光谱红移巡天项目, 覆盖数百万个星系、类星体和恒星。它利用重子声振荡 (BAO) 技术对暗能量进行测量, 通过构建宇宙关于演化时间的 3D 地图来研究暗能量对宇宙膨胀的影响。

为了给 DESI 项目选择合适的观测源, DESI Legacy Imaging Surveys, 即 DESI 遗珍成像巡天项目, 对 DESI 所包含的天区进行了测光巡天。DESI 遗珍成像巡天较为特别的一点是它由三个不同的子巡天项目构成, 分别为北京-亚利桑那天文巡天 (The Beijing-Arizona Sky Survey, BASS), 暗能量相机遗珍巡天 (The Dark Energy Camera Legacy Survey, DECaLS) 和 Mayall z 波段遗珍巡天 (The Mayall z-band Legacy Survey, MzLS)。其中, DECaLS 项目包含 g、r、z 三波段的数据, BASS 仅包含 g、r 波段的数据。

Euclid

欧几里得空间巡天望远镜由欧洲空间局发射, 包含 Y、J、H 三个波段的观测。它比目前最优秀的地基望远镜至少清晰四倍, 覆盖地球上共数十亿星系。欧几里得旨在绘制黑暗宇宙的几何形状并探索暗能量的奥秘, 通过观测弱透镜和星系团精确测量宇宙的膨胀历史和宇宙中大尺度结构的生长率。此项研究将使暗能量和修正引力区分开来, 进而探索宇宙加速的真正原因。目前欧几里得空间望远镜已经发射成功, 顺利开启探测任务, 并计划在 2025 释放第一批数据。

LSST

Large Synoptic Survey Telescope (LSST), 即大型综合巡天望远镜。它将利用

位于 Vera C. Rubin 天文台的 Simonyi 巡天望远镜，每隔几个晚上就遍历所计划的天区。因此，LSST 与其他巡天项目的不同点在于它能够检测瞬态天文事件，包括新星、超新星、伽马射线暴、类星体变率和引力透镜效应等等。除此之外，LSST 还计划通过测量弱引力透镜效应、重子声振荡和 Ia 型超新星的光度测定来研究暗能量和暗物质等。LSST 包含 u、g、r、i、z、y 六个波段的数据，可以观测非常暗弱的天体。单幅图像的完备星等 $M_r < 24.5$ ，在经过叠加的数据中可以达到完备星等 $M_r < 27.8$ 。

由于新冠疫情的影响，LSST 预计于 2025 年 8 月开始作业，并在两年后释放数据。

1.2.2 深度学习以及天文领域的应用

深度学习是机器学习的一个子集。与需要人工进行特征整理与提取的传统机器学习算法不同，深度学习是一种端到端的算法。在输入端（数据）与输出端（结果）之间，深度学习模型由特征层与输出层构成，实现无需任何其他步骤，由输入数据直接到输出结果。

深度学习的特征层由从浅到深的数个网络层构成，其目的是从输入的数据中提取相应的特征，形成“数据降维”。而在网络的输出层，则将相应的特征映射到所需的目的中，如分类、回归、图像生成等等。虽然深度学习的分类大相径庭，各自有各自专攻的网络类型与架构，但都逃不开特征提取与特征映射的过程。

尽管目前深度学习的广泛使用从结果论上证明了深度学习的有效性，但深度学习的发展历程非常艰辛。在很长一段时间内，由于性能无法比肩传统机器学习，深度学习一度无人问津。有正确的模型框架、优秀的数据集、强大的硬件算力等作为深度学习的硬门槛，从 1943 年 McCulloch 首次发布神经网络的雏形 MCP 模型以来，直到 2012 年，深度学习才借助 AlexNet 在 2012 年在 ImageNet 图像识别大赛中的一炮而红真正翻身。但从此以后，深度学习以图像领域与卷积神经网络为奇点遍地开花。任务与需求从简单的分类、回归拓展到了图像生成、目标识别，异常检测等复杂问题。网络模型本身也在卷积神经网络不断发展地同时 (Simonyan 等, 2014; He 等, 2016; Tan 等, 2019)，开发出了针对序列数据的 LSTM 网络 (Hochreiter 等, 1997)、针对图像生成的生成对抗网络 (GAN; Goodfellow 等, 2014)、变分自编码器 (Kingma 等, 2013, VAE:)、扩散模型 (Sohl-Dickstein 等, 2015)、针对目标识别的 YOLO 系列 (Wang 等, 2023) 等等。除此之外，还有不少模块如注意力机制 (Mnih 等, 2014) 被不断开发出来，用于提升不同应用场景下模型表现的性能。除了图像领域之外，自然语言处理与大模型如 ChatGPT、文心一言也是发展地如火如荼。

天文图像虽与传统图像相比，具有信噪比低、点扩散函数影响大、流量范围大、噪声背景区域多等诸多特点，但其作为图像的本质使得不少天文领域涉及图像的任务非常适合使用深度学习去解决。早在 2010 年，Banerji 就使用了神经网络

络去将斯隆巡天中的天体分类为早型星系、旋涡星系、恒星或人造卫星 (Banerji 等, 2010)。时至今日, 随着深度学习的兴起, 越来越多的天文工作者开始接触到深度学习与神经网络, 并思考合适的问题与需求。Xu 等 (2023) 使用了变分自编码器去提取星系的高维特征 (即隐变量), 再对隐变量进行离群值分析, 从而找到形态异常或被宇宙线、卫星等污染的星系图像。Euclid 空间望远镜的预研究中, 就使用 flow-VAE 从物理参数生成模拟的星系图像, 去仿真 Euclid 开始工作后所观测到的图像, 用于后续的程序测试 (Bretonnière 等, 2022)。除此之外, Kinakh 等 (2024) 还使用了超分从低分辨率的哈勃图像提升分辨率至 JWST 的图像。

1.3 利用不同方法的面亮度轮廓拟合程序

可靠且便捷、快速的面亮度轮廓拟合是许多大样本统计分析工作的基础。因此, 形形色色的面亮度轮廓拟合程序被大量的发布。围绕着共同的核心需求, 不同的程序根据不同的侧重点, 使用了各种不同的算法与代码结构, 适用于不同的硬件设备。根据上述的不同, 面亮度轮廓拟合程序可以大致分为传统方法、深度学习方法与 GPU 加速方法三种。目前, 传统方法仍是当前的主流算法。后面两种作为大数据时代的全新产物, 仍在不断发展当中。

1.3.1 传统方法

传统方法是最经典也是应用最广泛的拟合方法, 包括 GIM2D (Simard 等, 2002)、BUDDA (De Souza 等, 2004)、GALFIT (Peng 等, 2010)、GALFITM (Vika 等, 2013)、IMFIT (Erwin, 2015)、ProFit (Robotham 等, 2017) 等等。它们大部分诞生于编译语言盛行的时代 (如 C 语言、Fortran 等), 此时大样本巡天项目并不多见, 人们更多地使用它们进行小样本的精细拟合。因此, 使用传统方法的程序基本没有内置多样本同时输入的功能, 需要额外使用脚本进行自动化。同时, 编译语言的特性使得开发和维护具有挑战性, 用户进行程序的二次改写也较为困难。在算法上, 使用传统方法的程序基本使用 CPU, 利用不同的经典优化算法对面亮度轮廓拟合过程进行计算, 如蒙特卡洛马尔可夫链、嵌套采样、粒子群采样、差分进化、Levenberg-Marquardt (LM) 算法等等。下面, 我们介绍几个经典的使用传统方法的程序。

GALFIT

GALFIT 是面亮度轮廓领域最权威的程序之一, 首发于 2002 年 (Peng 等, 2002)。起初 GALFIT 只支持椭圆形态的模型, 但在其 2010 年释放的新版本 GALFIT3.0 中 (Peng 等, 2010), 新加入的傅里叶模式、弯曲、截断函数等支持了更详细的其他形态, 是目前内嵌轮廓种类最多的软件之一。除了基本版本之外, GALFIT 为 MEGAMORPH 特殊需求所开发的 GALFITM 支持多波段数据的联合拟合。而在未来的 GALFIT4.0 版本中, 将适配积分市场光谱仪所得到的 CUBE 数据。

GALFIT 程序由 C 语言进行编写, 其代码并不开源。因此, 用户无法自行对 GALFIT 的代码进行改写, 包括改变拟合优度、新增轮廓公式、改变最优化算法

等等。并且，GALFIT 作者表示，GALFIT 最初并不是用于自动批量拟合星系而设计，其在未来也不会有这样的计划。但尽管如此，目前已经有大量的管线路程序将 GALFIT/GALFITM 与目标检测的程序 SExtractor 连接在一起，实现自动批量的拟合星系，如 MEGAMORPH(Häußler 等, 2013), MORPHOFIT(Tortorelli 等, 2023), PYMORPH(Vikram 等, 2010) 等等。

在稳健性与速度之间，GALFIT 显然选择了速度。它所选择的 LM 算法是传统优化算法中最快的一种，但作为基于导数的算法的一种，其探索的参数空间较少，容易陷入局域最小值。目前，卡方统计是 GALFIT 唯一支持的拟合优度。

IMFIT

IMFIT 是由 Peter Erwin 于 2014 年首次发布 (Erwin, 2015)。作为开源软件，IMFIT 由 C++ 写就，并释放了基于 Python 的官方脚本，在天文已进入 Python 时代的当下使用方便。IMFIT 具备面向对象的设计，允许用户轻松加入新的轮廓公式。而在内嵌的轮廓中，IMFIT 除了常用的面亮度轮廓外，还包含有三维光密度模型的径向积分组件。IMFIT 允许用户在速度和稳健性之间进行权衡，可使用的优化算法包括 GALFIT 使用的 LM 算法（默认）、以及差分进化算法与单纯形法。与 LM 算法相比，后两个算法虽然速度较慢，但能遍历更多的参数空间，能减少被困于局域最小值的可能性。除了经典的卡方统计外，IMFIT 还支持基于泊松分布的最大似然统计，后者特别适用于低计数情况下的泊松数据。除此之外，IMFIT 还可以编译为使用多线程，利用现代 CPU 中的多个内核来提高速度。

PROFIT

PROFIT 是为 Galaxy And Mass Assembly (GAMA) 项目而开发，是基于贝叶斯的二维面亮度拟合程序，其代码完全开源 (Robotham 等, 2017)。PROFIT 中提供了多种轮廓模型，也支持用户对自己的轮廓模型进行自定义。与 GALFIT 等以梯度算法为核心的程序不同，PROFIT 在轮廓模型的形式上采取了积分的形式。实验证明，PROFIT 的积分算法能够比 GALFIT 更快、更准确地进行模型图像的计算。在优化算法与评价指标上，PROFIT 也提供了多种选择，如梯度算法、遗传算法、MCMC 方法进行优化；利用卡方分布或 Student-T 分布作为指标等等。在 PROFIT 中，参数可以选择在线性或对数空间中进行拟合。PROFIT 还支持自动绘制多种结果图像，除了常规的二维模型、残差图等，还支持将其投影至一维，并附有残差的分布图。

为了给 PROFIT 提供减去天空的星系流量图、观测误差图与初始参数，Robotham 等人还开发了 PROFIT 的配套软件 PROFOUND。其利用分水岭算法使得分割过程突破了椭圆形态的限制，令分割的结果更加准确 (Robotham 等, 2018)。

1.3.2 深度学习方法

利用深度学习方法的程序，其原理为利用神经网络建立图像的像素值到实数之间的回归映射。作为大数据时代与算力时代所激发的新方法，深度学习方法的主要目的是在准确性向传统方法靠拢的同时，速度大大提升。由于面亮度轮廓

拟合的本质为参数回归，因此，只需要准备好合适的训练集，面亮度轮廓拟合同样适合使用监督回归的神经网络。由于输入的是星系图像，本质上这可以归类为计算机视觉类的工作，使用卷积神经网络将会是比较常规的做法。

目前，使用深度学习方法进行面亮度轮廓拟合的工作仍在发展当中，相关的工作数量较少。第一个使用深度学习预测轮廓参数的工作为 Tuccillo 等 (2018) 发布的卷积神经网络模型 DeepLeGATo。首先，他利用 GALSIM 产生的模拟星系，并叠加噪声之后组成的数据集训练了可预测单 Sérsic 面亮度轮廓的卷积神经网络。随后，他使用 GALFIT 与 DeepLeGATo 同时对测试集（没有参与模型训练的数据）中的星系进行轮廓拟合/参数预测，并证明 DeepLeGATo 在准确性上与 GALFIT 相当的同时，速度提高了约 50 倍。接下来，Tuccillo 还将 DeepLeGATo 应用在了 CANDLES 数据上，其结果起初并不理想，但在进行迁移学习后得到了大幅的改善。然而，这种迁移治标不治本，当运用到另一个天区中时，此方法依然需要进一步的迁移。

与 DeepLeGATo 非常相似，Ghosh 等 (2022) 通过 GALFIT 制作的模拟星系作为训练集，并证明通过迁移学习，所训练的卷积神经网络 GAMORNET 可以为低红移星系和高红移星系提供可靠的轮廓参数，并将其运用在 SDSS 与 CANDELS 巡天约 120,000 个星系上，进行了关于恒星形成的研究。Li 等 (2022) 利用纳入/不纳入真实点扩散函数图像的模拟星系训练集，并发现当点扩散函数纳入训练的时候，其卷积神经网络 GalNet 无需迁移学习即可获得很好地结果，且速度快了 3 倍。作为 GalNet 的后续研究，Qiu 等 (2023) 等使用 GALSIM 的模拟数据集，并加叠上 CSST 的模拟背景场，训练了可预测盘 + 核球双成分的卷积神经网络，并进行了 B/T 的参数验证，证实了其网络用于成分分解的可行性。同样，Grover 等 (2021)，也使用了卷积神经网络进行 B/T 的预测，并表明可以在不到一分钟的时间内估计 20,000 个星系的核球与盘之间的光度比。

在这些深度学习的方法中，网络模型在星系图像和轮廓参数之间建立了端到端的新映射，从而避免了对人工的依赖，并大幅度提升了计算轮廓参数的速度。但深度学习方法的弊端在于稳健性不强。端到端的新映射是一种“黑盒子”映射，不具备较好地可解释性，因此在某种巡天或模拟数据下训练的深度学习网络，无法很好地适应到其他不同巡天的数据。尤其是地基与空基之间，参数空间的域相差非常大。如何实现模型的外推成为了值得考量的问题。纳入不确定性估计是其中一种可能的解决方式。Ghosh 等 (2022) 使用了空间注意力机制网络与蒙特卡洛 dropout 的组合对预测的不确定性进行估计。除此之外，Aragon-Calvo 等 (2020) 使用了自监督混合网络架构，以避免使用基于模拟的完全监督训练。

1.3.3 GPU 加速方法

在进入算力时代的当下，若要避免使用深度学习导致的“黑盒子”问题，我们可以利用 GPU 的强大算力与自动求导帮助我们加速传统方法。GPU 可以帮助加速由参数到星系图像的过程，而自动微分可以避免显示求导，在优化过程中

自动计算导数，便于实现面向对象的程序。自动微分方法与有限差分方法计算导数不同 (Baydin 等, 2018)，自动微分方法的计算成本不会随着参数的数量而变化，且精度非常高。目前，许多高度集成与开发的深度学习框架都支持 GPU 加速与自动微分。它们在底层使用 CUDA 语言编写，从而避免了天文工作者在较为复杂与困难的 CUDA 编程中消耗不必要的时间。除此之外，它们还对自动微分方法进行了许多优化与高性能的视线，如 PyTorch 所支持的动态梯度图等。

目前，GPU 加速方法的应用非常广泛。Modi 等 (2021) 开发了名为 Mesh-TensorFlow 的程序，利用 GPU 加速、分布式和可微分模拟以实现粒子网格 (PM) 宇宙学 N 体代码，在大型超级计算机上。作者发现，利用了 GPU 加速之后，比原先基于 Python 的代码快了十倍。类似的，Schoenholz 等 (2020) 开发了名为 JAX MD 的软件包，依托 JAX 框架进行分子动力学可微分物理模拟。由星系动力学建模到宇宙 N 体模拟。在星系面亮度轮廓领域，也逐渐开始有程序开始使用 GPU 加速与自动微分。BANG (Rigamonti 等, 2023) 和 PYSERSIC (Pasha, 2022) 使用了 GPU 加速进行星系模型图像的计算，但它们并未采用自动微分，而是使用了 NUTS 采样器与嵌套采样。最近，Stone 等 (2023) 开发的 AstroPhot 是面亮度轮廓领域中首先结合 GPU 加速与自动微分进行面亮度轮廓拟合的程序。然而，AstroPhot 更多地面向多波段联合拟合的，而非进行大样本星系的自动拟合。其 GPU 加速针对于单张大像素的星系模型图像的生成。

使用 GPU 加速与自动微分方法加速传统拟合方法，不仅可以快速进行面亮度轮廓拟合以适应大数据时代，还非常好的保留了物理的可解释性。

1.4 本文概览

1.4.1 本文研究目标

综上所述，大样本巡天时代将有机会为星系形成与演化的研究做出全新的贡献。为了更好地利用起未来将面临的海量数据，我们预期开发一套新的面亮度轮廓拟合程序 GALMOSS。

在所设想的 GALMOSS 工作流程中，我们将大样本数据的并行化拟合纳入程序设计目的，将批量的星系图像数据集处理为多个多维数组。在正向的星系模型图像生成过程中，所有的计算过程将被抽象为多维数组的计算，利用 PyTorch 深度学习框架实现 GPU 加速与自动微分。这种基于张量的方法允许通过调整批量大小和优化 GPU 利用率来动态管理计算期间的数据大小。处于对速度、计算代价以及大样本拟合目标的考虑，在图像的拟合过程中，GALMOSS 将使用梯度下降进行参数的优化，因为梯度下降的计算和内存要求相对较低。为了解决梯度下降不自然包含参数拟合误差估计的问题，我们还将程序中提供了协方差矩阵与自举重采样两种不确定性估计方法。我们预期 GALMOSS 将应当继承传统方法的可解释性与现代方法的强大算力，兼具拟合准确性与快速的特征。

1.4.2 本文结构

本文的主要内容如下。在第2节我们主要介绍了 GALMOSS 程序的工作流程，并在第3节主要介绍了对 GALMOSS 性能的实验。在第4节，我们对本文的工作进行总结，并对未来博士期间的工作进行展望与规划。

第2章 GALMOSS: 基于 GPU 加速的星系面亮度拟合程序

如果用一句话来概括 GALMOSS 的核心, 那就是使用 PyTorch 作为一种“编程语言”, 按照传统面亮度拟合程序的思路编写, 但为大批量并行化拟合特殊设计的新程序。因此, GALMOSS 将拥有速度快、用户友好、拟合结果可解释性强等特点。在本章节, 我们将对 GALMOSS 的这些创新性特点进行说明。首先, 我们对 GALMOSS 程序整体的工作流程做一个总体的介绍; 其次, 我们将介绍 GALMOSS 的前向建模过程以及如何针对此过程进行加速; 再次, 我们将介绍 GALMOSS 的参数优化方法; 最后, 我们将介绍 GALMOSS 如何对拟合误差进行估计。

2.1 工作流程

如图2.6, GALMOSS 首先进行的是读取数据。GALMOSS 的正常运行需要用户定义的初始参数值 p , 并提供由数据图像 D 、观测误差图像 σ (描述观测数据的不确定性)、掩膜图像 Ω 与点扩散函数图像 PSF 组成的图像数据集。用户定义的初始参数值可以分为几何参数集 p_G 与轮廓参数集 p_L 两类。几何参数集 p_G 描述了星系轮廓模型的几何属性, 即定义了椭圆等照度线的中心位置、轴比与方位角。轮廓参数集 p_L 则定义了星系轮廓模型径向亮度的初始值 $\Phi(r)$, 即表示了亮度将如何随半径 r 变化。

GALMOSS 获取这些初始值后, 将根据用户所选定的轮廓模型种类生成相对应的原始模型图像。为了模拟观测效应, 原始模型图像需要与对应星系图像观测时所属的点扩散函数图像进行卷积。当有多个模型用于拟合一张星系图像时, 点扩散函数图像将与累加后的总模型图像进行卷积。随后, 若天空背景指定与其他模型一起参与星系图像拟合, 将在卷积过程完成后直接逐像素添加平均天空流量。由于天空背景的观测效应基本可以忽略, 点扩散函数图像无需与天空流量进行卷积。如此, 最终的图像模型 $M_{i,j}$ 就完成了完整的建模过程。这一过程利用并行化的计算设置得到 GPU 的充分加速。

接下来, 为了优化模型参数 p_G 与 p_L , 图像模型 $D_{i,j}$, 观测误差图像 $\sigma_{i,j}$ 将共同参与优化指标 χ^2 的计算。与此同时, 掩膜图像 $\Omega_{i,j}$ 的加入将去除可能造成负面影响的像素 (如前景恒星等), 使得优化的过程更加稳健。在得到优化指标后, GALMOSS 将利用 PyTorch 框架中提供的自动反向传导实现梯度下降的优化过程。这一过程同样得到 GPU 强大并行能力的支持。

达到所设置的优化次数或提前达到优化目标后, 拟合过程完成。GALMOSS 将利用最优的拟合结果进行拟合误差的估计。随后, GALMOSS 会将数据产出保存到用户指定的输出路径。数据产出包括图像数据与拟合参数两种。图像数据以 FITS 格式, 且以单个星系为单位进行分别保存 (Img_block)。在图像数据中, 将

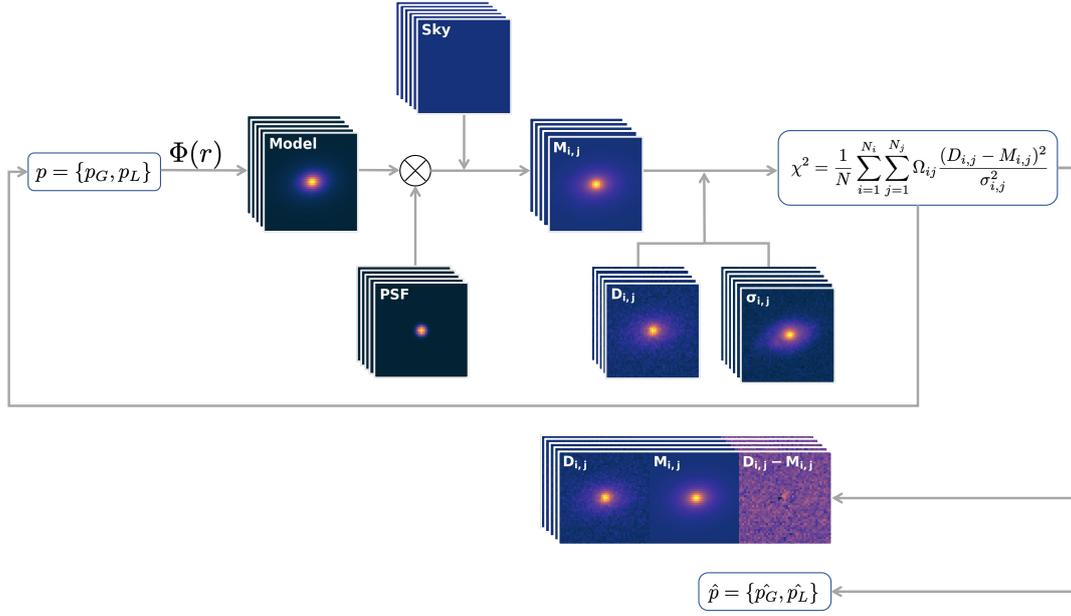


图 2.1 GALMOSS 的工作流程：从参数初始化和图像处理到拟合过程和数据产出的导出。最终的数据产出包括拟合模型和残差，默认全部以 FITS 格式存储。

包含有对应星系的数据图像、总模型图像、残差图像以及子成分图像（如设定多个轮廓模型同时拟合）。拟合参数包含同时拟合的所有星系的 χ_μ 和模型参数 p ，可以选择以 FITS、hdf5、csv 三种格式保存。

2.2 图像的生成

对于给定的一组星系或轮廓参数，GALMOSS 根据所选的面亮度轮廓模型生成表示每个像素的强度或表面亮度的模型图像。GALMOSS 内置了七种常用的径向面亮度轮廓，分别为 Sérsic、Exponential disk、Ferrer、King、Gaussian、Moffat，以及 Flat sky 模型。这些模型无需用户再定义，可以直接被调用。而并未被内置的面亮度轮廓，也可以非常方便地利用 GALMOSS 进行定义。除了使用单面亮度轮廓外，GALMOSS 还支持选择多个轮廓进行组合拟合。

在轮廓被选定后，这些径向轮廓将利用椭圆的对称性被转换为二维轮廓并生成图像，便于后续的优化指标计算。接下来，我们将考虑一个简单的使用单面亮度轮廓进行投影的过程来更具体地介绍图像生成的过程。考虑到 Sérsic 轮廓在星系面亮度拟合研究中的广泛应用，我们将使用 Sersic 轮廓为例，并随后介绍其他面亮度轮廓。

2.2.1 从参数到模型图像

Sersic 轮廓

Sérsic 轮廓，又称 $r^{1/n}$ 律，通常用于拟合椭圆星系或其他星系类型的盘或者

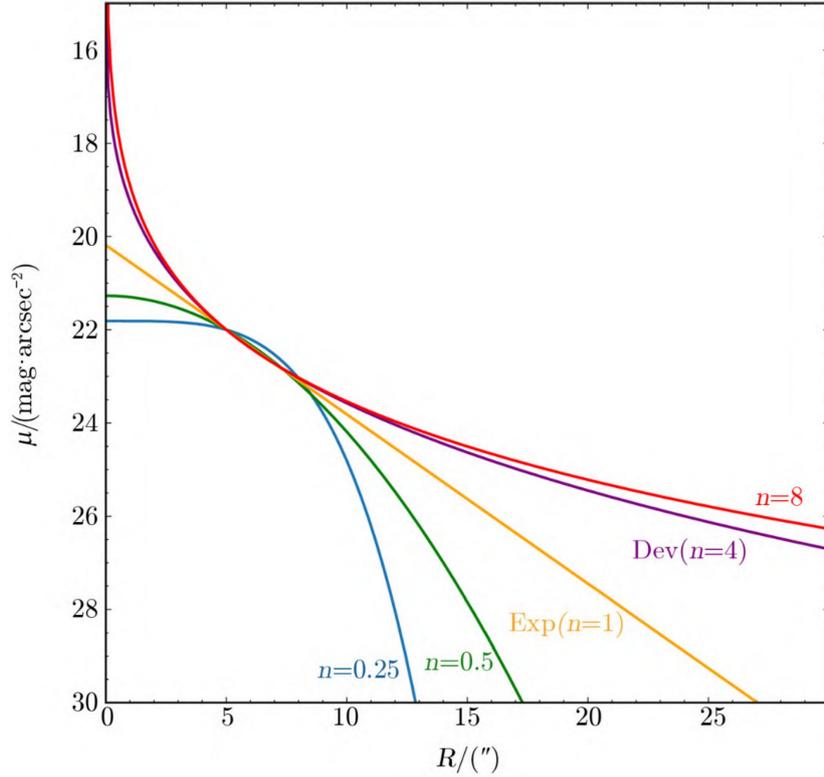


图 2.2 Sérsic 轮廓随指数因子 n 的变化。图源 Shen 等 (2022)。

核球部分的表面亮度轮廓。作为半径 r 的函数，其光强 $I(r)$ 可以由下式进行定义：

$$I(r) = I_e \exp \left\{ -v_n \left[\left(\frac{r}{r_e} \right)^{\frac{1}{n}} - 1 \right] \right\}, \quad (2.1)$$

其中， I_e 表示有效半径 r_e 处的表面亮度，而有效半径 r_e 则定义为包含总轮廓一半流量时其半径的位置。

Sérsic 指数 n 表示了面亮度轮廓变化的趋势。参数 v_n 并非自由参数，是 Sersic 指数 n 通过伽玛分布的逆累积分布函数进行数值计算而得。Sérsic 在诸多类似的面亮度轮廓中并不是最早出现的，但却是最有概括性与灵活性的。通过调整 Sérsic 指数 n 的数值，Sersic 轮廓可以模拟各种星系形态。例如，设置 $n = 4$ 会将其转换为 de Vaucouleurs 轮廓，即 $r^{1/4}$ 律。当 $n = 1$ 时，Sersic 轮廓会转变为 Exponential Disk 轮廓，而 $n = 0.5$ 产生高斯分布轮廓。只需要简单地改变 n ，用户即可探索这些不同的轮廓。

面亮度轮廓一般以一维径向的方式来定义（如式2.1），而半径 r 指的是当前像素距星系中心的距离。但当面亮度轮廓被转换到二维方式时，半径 r 就不再是简单的距中心距离，而是当前像素所对应的椭圆等效半径。当给定椭圆中心 (x_c, y_c) 、椭圆的方位角 θ 和轴比 q 后，椭圆的等效半径 r 定义如下（参考 Robotham 等, 2017）：

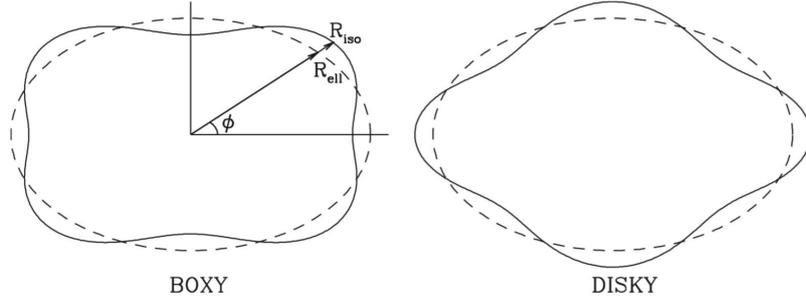


图 2.3 盒状星系与盘状星系。

$$r = \left[r_{\text{maj}}^{(B+2)} + \left(\frac{r_{\text{min}}}{q} \right)^{(B+2)} \right]^{\frac{1}{B+2}}, \quad (2.2)$$

其中,

$$r_{\text{maj}} = |\cos(\theta)(x - x_c)| + |\sin(\theta)(y - y_c)|, \quad (2.3)$$

$$r_{\text{min}} = |-\sin(\theta)(x - x_c)| + |\cos(\theta)(y - y_c)|. \quad (2.4)$$

盒状参数 B 对椭圆的形状进行了拓展 (Athanasoula 等, 1990), 增加了椭圆等面亮度线形状的灵活性, 使其可以呈现不同的矩形形态。它将标准椭圆拓展为远离标准 L_2 范数的菱形外观 (L_1 范数) 或矩形外观 (L_∞ 范数)。如图 2.3 所示, $B = 0$ 的标准椭圆如虚线所示; 当 $B > 0$ 时, 椭圆将如图左所示, 更多的呈现矩形形态, 称为盒状。盒状在长轴、短轴的位置向上, 等效半径将比标准椭圆要小; 而在长轴短轴之间的位置, 等效半径将超出标准椭圆的范畴。当 $B < 0$ 时, 椭圆将向盒状星系相反的方向变化。如图右所示, 椭圆的形态如菱形一般, 称为盘状。

当等效半径定义完毕后, 理论上 GALMOSS 已经可以直接由轮廓公式出发, 进行从参数到模型图像的计算了。然而, 在面亮度轮廓拟合的实际应用中, 直接使用 Sérsic 轮廓的有效光强 I_e 在拟合中提供流量水平的约束并不是常用的做法。更广泛的用法是使用星等 m 及其相应的星等零点 m_0 来约束该面亮度轮廓的总流量。这种方法更便于与观测数据进行比较, 是许多被广泛使用的星系面亮度轮廓拟合程序默认的设置 (例如 GALFIT (Peng 等, 2010) 和 PROFIT (Robotham 等, 2017))。为了从有效光强 I_e 转向使用星等 m , 两个参数之间的转换关系可以由数值方法定义如下:

$$I_e = \frac{r_{\text{box}} v_n^{2n} 10^{-0.4(m-m_0)}}{2\pi q r_e^2 n \Gamma(2n) \exp(v_n)}, \quad (2.5)$$

其中

$$r_{\text{box}} = \frac{\pi(2+B)}{2\beta \left(\frac{1}{2+B}, \frac{1}{2+B} \right)}, \quad (2.6)$$

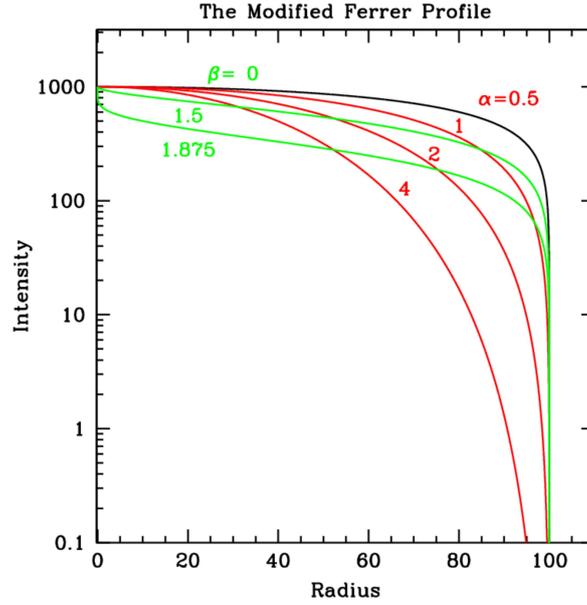


图 2.4 修正的 Ferrer 随斜率参数的变化, 图源 Peng (2003)。

位于分母位置的 beta 函数 $\beta(a, b)$ 使用以下关系计算 $\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, Γ 表示 Gamma 函数。在这个转换关系中, r_{box} 为 I_e 的几何校正因子, 以修正当等效半径 r 被盒状参数修改后, 偏离了完美椭圆而对星等-流量关系转换关系严格性的破坏。比如, 当椭圆保持原有形态时, $B = 0$ 时, 因而 $r_{\text{box}} = 1$, 对转换关系没有影响, 即表示没有几何校正。而当 $B \neq 0$ 时, $r_{\text{box}} \neq 1$, 几何修正将在转换关系中起到一定作用。

除了天空背景轮廓之外, 其他 GALMOSS 的内置轮廓均径向轮廓, 适用公式 2.2 到 2.4。接下来, 我们对其他内置轮廓进行简要描述。

修正的 Ferrer 轮廓

修正的 Ferrer 轮廓光强分布的特点为核心接近平坦, 而到了轮廓外围流量迅速降低并截断。Ferrer 轮廓最初由 Ferrers (1877) 提出, 后来由? 进行修正。一个服从修正 Ferr 轮廓的模型, 其光强随半径的变化可以定义如下:

$$I(r) = I_0 \left[1 - \left(\frac{r}{r_{\text{out}}} \right)^{2-b} \right]^a, \quad (2.7)$$

其中 I_0 为轮廓中心的面亮度, r_{out} 为外截断半径, b 和 a 分别是控制光强截断部分和核心部分斜率的轮廓参数。与 Sersic 轮廓类似, GALMOSS 在实际运用中使用星等 m 及其相应的星等零点 m_0 来约束该面亮度轮廓的总流量。其换算关系定义如下:

$$I_0 = \frac{r_{\text{box}} 10^{-0.4(m-m_0)}}{\pi q r_{\text{out}}^2 a \beta \left(a, 1 + \frac{2}{2-b} \right)}, \quad (2.8)$$

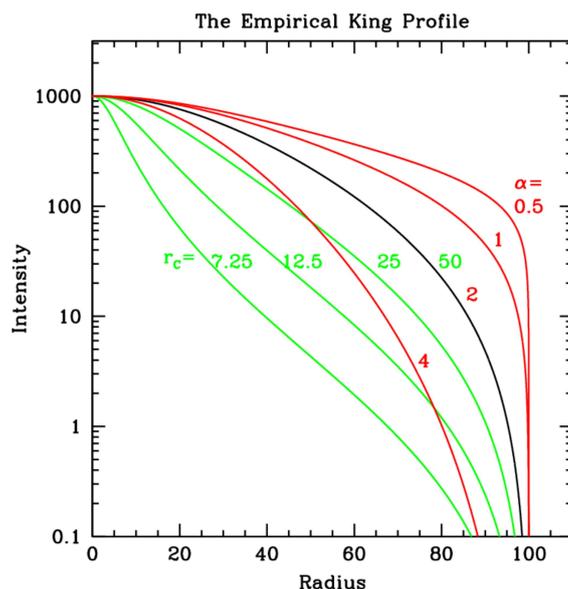


图 2.5 King 轮廓随核半径、幂律因子的变化，图源Peng (2003)。

其中 q 为星系轴比， β 与式2.6中所使用的 β 形式相同。

King 轮廓

King 轮廓常用于拟合银河系内与河外星系中的球状星团。一个服从 King 轮廓的模型，其光强随半径的变化可以定义如下：

$$I(r) = I_0 \left[1 - \frac{1}{\left(1 + \left(\frac{r_t}{r_c}\right)^2\right)^{1/\alpha}} \right]^{-\alpha} \times \left[\frac{1}{\left(1 + \left(\frac{r_t}{r_c}\right)^2\right)^{1/\alpha}} - \frac{1}{\left(1 + \left(\frac{r}{r_c}\right)^2\right)^{1/\alpha}} \right]^{\alpha} \quad (2.9)$$

式中， I_0 为轮廓中心的面亮度。核半径 r_c 表示密度开始偏离均匀性的尺度，而截断半径 r_t 则代表了星团的边界。球状星团幂律因子 α 决定表示了密度随着距中心距离的增加而下降的速率。球状星团中恒星的浓度可以使用参数 r_t 和 r_c 定义，表示为 $c = \log\left(\frac{r_t}{r_c}\right)$ 。

moffat 轮廓

Moffat 轮廓常用于拟合一个真实望远镜的点扩散函数。一个服从 Moffat 轮廓的模型，其光强随半径的变化可以定义如下：

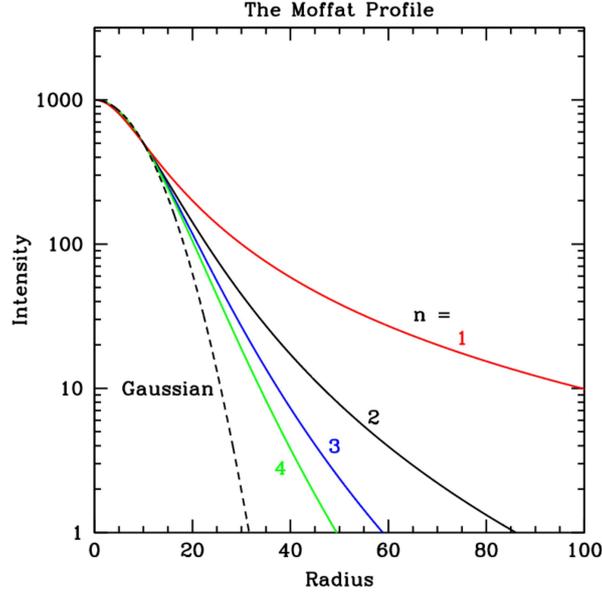


图 2.6 Moffat 轮廓随指数因子的变化，图源Peng (2003)。

$$I(r) = I_0 \left[1 + \left(\frac{r}{r_d} \right)^2 \right]^{-n}, \quad (2.10)$$

其中

$$r_d = \frac{\text{FWHM}}{2\sqrt{2^{\frac{1}{n}} - 1}}. \quad (2.11)$$

式中 I_0 为轮廓中心的面亮度。聚集度指数 n 指示分布是近似于洛伦兹分布 ($n=1$) 或近似于高斯分布 ($n \rightarrow \infty$)。与 Sersic 轮廓相似，GALMOSS 在实际运用中使用星等 m 及其相应的星等零点 m_0 来约束该面亮度轮廓的总流量，其换算关系定义如下：

$$I_0 = \frac{r_{\text{box}}(n-1)10^{-0.4(m-m_0)}}{\pi q r_d^2}. \quad (2.12)$$

其中 q 为轴比, r_{box} 的定义与2.6相同。

指数盘轮廓

一个服从指数盘轮廓的模型，其光强随半径的变化可以定义如下：

$$I(r) = I_0 \exp\left(-\frac{r}{r_s}\right), \quad (2.13)$$

其中 I_0 为该轮廓在中心的面亮度（即半径 $r = 0$ ），而 r_s 为指数盘的标长。与 Sersic 轮廓类似，GALMOSS 在实际运用中使用星等 m 及其相应的星等零点 m_0 来约束该面亮度轮廓的总流量。给定标长 r_s 后，其换算关系定义如下：

$$I_0 = \frac{r_{\text{box}} 10^{-0.4(m-m_0)}}{2\pi q r_s^2}, \quad (2.14)$$

其中 q 为轴比, r_{box} 的定义与2.6相同。

高斯轮廓

高斯分布的经典应用包括对简单的点扩散函数或点源进行建模。一个服从高斯轮廓的模型, 其光强随半径的变化可以定义如下:

$$I(r) = I_0 \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad (2.15)$$

其中 I_0 为该轮廓在中心的面亮度 (即半径 $r = 0$), σ 为径向色散。尽管高斯轮廓由 σ 定义, 但 GALMOSS 在实际运用中使用半高全宽 FWHM 进行拟合。 σ 与 FWHM 的换算关系定义如下:

$$\text{FWHM} = 2.354\sigma, \quad (2.16)$$

天空轮廓

与其他径向轮廓不同, 在 GALMOSS 中, 天空轮廓为一个处处相等的平均天空流量, 与所在的半径无关:

$$I = I_{\text{sky}}. \quad (2.17)$$

在 GALMOSS 中, 天空轮廓唯一需要的参数为平均天空流量值 I_{sky} 。如果用户需要其他更复杂的天空轮廓, 可使用 GALMOSS 的用户自定义功能。

组合与自定义轮廓

复杂星系的面亮度轮廓无法用单一的轮廓来描述。相类似, 若星系图像由多个天体叠加而成, 也同样需要多个轮廓同时进行拟合。在 GALMOSS 中, 只要运行内存支持, 可以选择任意数量的轮廓进行拟合。如在进行核球 + 盘的面亮度分解时, 仅需定义两个 Sersic 轮廓, 叠加后同时与星系图像进行拟合。除此之外, GALMOSS 作为自由度高、用户友好度高的拟合程序, 也支持自定义面亮度轮廓。用户只需要完成正向的建模过程即可。

2.2.2 PSF 卷积

在上一章节, 我们已经可以通过几何参数 p_G 、轮廓参数 p_L 进行基于所选轮廓的星系模型图像计算。然而, 此时的星系模型图像还不够“真实”。真实的星系流量在被 CCD 接收时, 因为视宁度、探测器的光电子扩散等原因, 流量分布将出现一定的偏差。这一效应将在下一章节中通过卷积点扩散函数图像进行解决。

点扩散函数的主要来源根据望远镜的仪器与所属环境有很大关系。如地基望远镜的点扩散函数由视宁度占主导，但在天基望远镜中，大气湍流已经微乎其微，因此天基望远镜的点扩散函数远优于地基望远镜。在点扩散函数的获取中，较为精确的方法是使用多个恒星或类星体的观测图像去表现实际的点扩散效应；而较为简单的方法则是给定半高全宽，利用高斯分布来近似。

为了模拟流量的偏差，点扩散函数需要与模型图像进行卷积。在卷积的过程中，利用空间域的直接卷积与频域的 FFT 卷积是两种常用的卷积方法。在直接卷积的过程中，点扩散函数将以滑动窗口的方式遍历整张图像，并进行逐像素的计算。而在 FFT 卷积的过程中，模型图像与点扩散函数图像都将利用傅里叶变换转到频域，在频域中进行逐元素的计算，再使用逆 FFT 将结果转回空间域。在点扩散函数相对较小的时候，FFT 卷积与直接卷积相比速度较慢。但随着模型图像、PSF 图像尺寸的增长，直接卷积的时间增长将与图像尺寸线性相关；而频域卷积的时间复杂度则 $\sim n \log n$ ；当点扩散函数图像大小与模型图像大小相当时，使用频域卷积将更为合适；然而，频域卷积的计算开销较大，对内存的需求量较高（详见 Astropy 的对两种方法的测试¹）。而且，这两种方法在精度上并没有显著差别 (Robotham 等, 2017)。因此，考虑到 GALMOSS 并行化拟合的需求，程序需要尽量减少对运行内存的占用，我们选择在 GALMOSS 中使用空间直接卷积。

在 GALMOSS 中，星系模型图像的并行卷积操作使用由 PyTorch 框架提供的直接卷积方法。在对输入的点扩散函数进行归一化后，PyTorch 将对模型图像外围进行自动补零，使星系图像经过卷积后仍然维持相同的大小。由于 PyTorch 内置的卷积操作并不包含对角转换，GALMOSS 会对点扩散函数进行对角转换后再进行卷积操作。

2.2.3 GPU: 图像生成过程的加速原理

前文介绍了 GALMOSS 完整的的图像生成过程。若将这一过程抽象为向量与矩阵的乘法，则不论是几何参数到椭圆等面亮度线、或者是椭圆带入径向面亮度轮廓、亦或者是点扩散函数进行卷积，都可以视为矩阵乘法及其组合来解决。当多个模型图像进行并行化生成时，二维矩阵则变为高维矩阵，其矩阵乘法的本质并未得到改变。而在深度学习时代，GPU 的核心竞争力之一，就是能够快速进行（多维的）矩阵乘法。

矩阵乘法本身即具有大量互不相关的计算过程，这是 GPU 可以进行矩阵乘法运算加速的基本原理。如图2.7所示的矩阵乘法过程，是大小为 M, K 的矩阵 A 与大小为 K, N 的矩阵 B 相乘后得到大小为 M, N 的矩阵 C 。当这一计算过程在 GPU 中进行加速时，矩阵 C 将被裁剪为不同的子块 C' ，其大小为 M_{tile}, N_{tile} 。在每一个 C' 的计算过程中，其计算的线程将读取矩阵 A 、矩阵 B 相应的子块 A' 与 B' ，并逐步遍历 K 维度，反复读取、相乘、累加，直到完成矩阵 C' 的计算。

¹https://keflavich-astropy.readthedocs.io/en/convolve_fft_profiling/convolution/performance.html

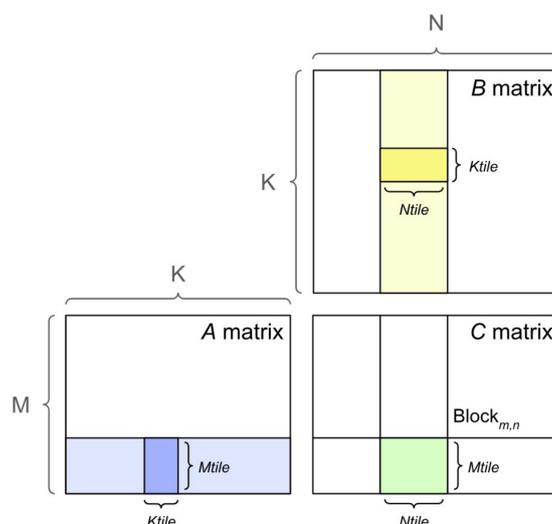


图 2.7 GPU 矩阵计算加速的原理，图源<https://docs.nvidia.com/>。

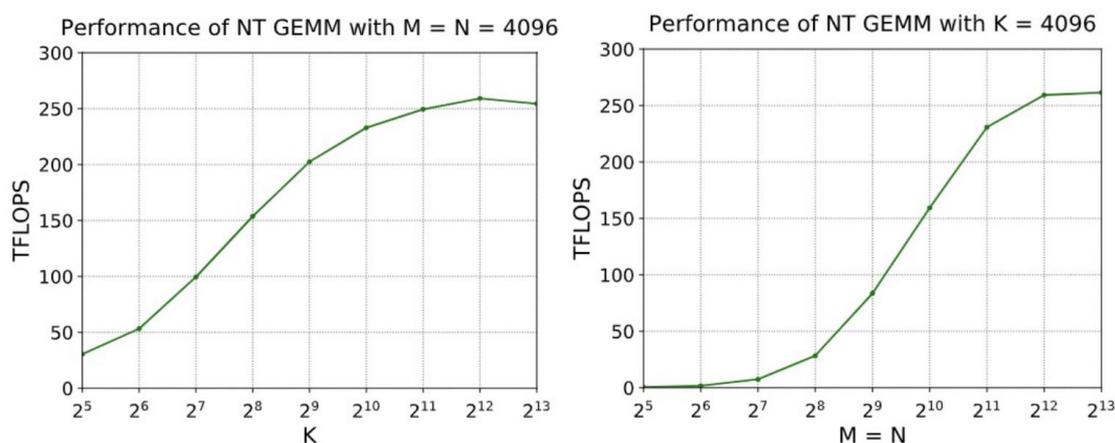


图 2.8 GPU 矩阵计算对算力的使用，图源<https://docs.nvidia.com/>。

因此，不论多小的矩阵，都能在 GPU 中得到一定的加速。

但使用 GPU 加速小矩阵，并不能充分利用起 GPU 的强大算力。FLOPS，即每秒浮点运算次数，是衡量 GPU 算力的重要指标之一。图2.8表示了不同大小的矩阵对 GPU 算力应用的情况。左图表示了当矩阵的 M, N 恒为 4096 时，矩阵计算的 FLOPS 随 K 的增长而增长，在 2^{12} 左右达到饱和；右图表示了当矩阵的 K 恒为 4096 时，矩阵计算的 FLOPS 随 M/N 的增长而增长，在 2^{13} 左右达到饱和。右图可知，数据量大的矩阵运算才能充分调动 GPU，实现最大程度的加速。

在大样本的前提下，GALMOSS 并行生成模型图像的计算量足以达到充分调动 GPU 的标准。充分利用 GPU 的性能对模型图像的生成过程进行加速，这也是 GALMOSS 快速拟合性能的重要落脚点之一。

2.3 图像的拟合

2.3.1 拟合优度

当数据服从正态分布时, 非常适合使用卡方分布与卡方统计量进行拟合优度检验。由于天文学数据经常使用正态分布来近似, 卡方统计被广泛用做天文学模型拟合过程中的拟合优度指标。

天文观测图像也可以近似为正态分布。当 CCD 接收到遥远天体发出的光信号时, 光信号的流量服从泊松分布。而当接收到的光信号足够大时, 泊松分布可以近似视为正态分布。因此, 如果模型图像能很好地还原观测图像的流量, 那么模型图像与观测图像之间的差异应该仅由观测图像的观测误差贡献。根据以上推断, 我们能够建立一组满足服从标准正态分布的数据:

$$\frac{D_{1,1} - M_{1,1}}{\sigma_{1,1}}, \dots, \frac{D_{1,j} - M_{1,j}}{\sigma_{1,j}}, \dots, \frac{D_{i,j} - M_{i,j}}{\sigma_{i,j}}, \quad (2.18)$$

并定义

$$\chi^2 = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \Omega_{i,j} \frac{(D_{i,j} - M_{i,j})^2}{\sigma_{i,j}^2}. \quad (2.19)$$

由卡方分布的定义我们得知, 若一组数据为独立的标准正态随机变量, 则它们的平方和满足自由度为 k 的卡方分布。且当自由度 k 足够大时, $f(x; k)$ 将会逼近均值为 k , 方差为 $2k$ 的正态分布。

根据卡方分布的分布函数, GALMOSS 的对数似然函数为:

$$\ln \mathcal{L} = -\frac{k}{2} \ln 2 - \ln \Gamma(k/2) + \left(\frac{k}{2} - 1\right) \ln \chi^2 - \frac{\chi^2}{2}, \quad (2.20)$$

其中自由度 $k = N - m$, 是未包含在掩模中的像素数量 N 与模型拟合参数数量 m 之间的差。在模型与观测数据符合良好, 即最大化似然函数时, χ^2 应当趋近于正态分布的均值 k 。因此, 我们定义每自由度的卡方为:

$$\chi_v^2 = \frac{\chi^2}{v}, \quad (2.21)$$

χ_v^2 的数值在理想的情况下应当趋向于 1, 当 $\chi_v^2 \gg 1$, 则说明模型拟合不佳, 而当 $\chi_v^2 < 1$ 则说明模型过拟合。

2.3.2 参数优化方法

在已经得到拟合优度的定义后, 如何优化拟合优度则是“各显神通”了。在被广泛应用的面亮度拟合程序中, 除了 IMFIT(Erwin, 2015) 提供的差分进化算法 (Differential Evolution; Storn 等, 1997), AutoGalaxy(Nightingale 等, 2023) 提供的嵌套采样 (Nested sampling; Speagle, 2020)、粒子群采样方法 (Particle Swarm

Optimization; Miranda, 2018), 以及 IMFIT 和 BUDDA(De Souza 等, 2004) 均可使用的无约束单纯形方法 (Nelder-Mead simplex method; Nelder 等, 1965) 之外, 目前常用的拟合算法可以被大致分为马尔可夫链蒙特卡罗方法与梯度方法两种。

马尔可夫链蒙特卡罗方法

马尔可夫链蒙特卡罗方法, 又简称 MCMC 方法, 是由马尔可夫链和蒙特卡罗两种数学方法所共同组成。蒙特卡罗作为一种计算方法, 普遍应用于求解难以直接计算的数值结果 (Hammersley, 2013; Rubinstein 等, 2016)。其通过构造模型, 使得模型采样结果的某些统计性质与所求的数值结果相等, 从而巧妙地解决求解困难的问题。譬如一个相切于正方形的圆形的经典模型, 它可以通过随机撒点的结果来估计难以直接计算的圆形面积。但当模型的分布非常复杂的时候 (如高维情况), 直接采样往往会遇到维数灾难, 而这可以使用马尔可夫链的性质来解决。马尔可夫链是描述系列可能事件的随机模型, 其每个事件的概率 P_i 只与前一个事件的状态有关 (Meyn 等, 1994)。换言之, 马尔可夫链是遵循某种规律 (转移矩阵) 从起始点开始往后延展的。当一个马尔可夫链满足收敛条件时 (Cowles 等, 1996), 最后事件的分布必将收敛到某一个唯一的平稳分布, 且起始点的位置并不影响其结果。当此唯一平稳分布等于所需要采样的复杂分布时, 收敛后的事件的集合即等于我们所需要的采样结果。当运用在拟合问题上时, 所求的模型参数与拟合的不确定度等于马尔可夫链平稳分布的期望与方差。

MCMC 方法拟合分布的本质使得它能更广泛地搜索参数空间, 且其拟合的不确定度是由马尔可夫链的方差所天然得出。这一优势使得 MCMC 方法被广泛应用于天文领域的拟合问题 (PROFIT; Robotham 等, 2017)。纵然 MCMC 方法已经减少了直接采样的困难程度, 但其巨大的计算量依然是难以解决的问题。因此, 一系列更复杂、性能更优秀的 MCMC 方法逐渐应用于包括星系面亮度轮廓的拟合的天文领域。GIM2D(Simard, 1998) 使用了 Metropolis 算法 (Hastings, 1970; Metropolis 等, 1953), 其改进了 MCMC 方法接受率过低的问题, 加快了收敛的速度。AutoGalaxy 使用了 MCMC 采样器 emcee 包 (Foreman-Mackey 等, 2013), 在处理高维问题时能加速收敛。

基于导数的优化算法

基于导数的优化算法有多个变种, 如梯度下降法 (Cauchy 等, 1847)、牛顿法、高斯-牛顿法 (Gauss, 1877)、LM 算法 (Levenberg-marquardt; Levenberg, 1944; Marquardt, 1963) 等等。它们均使用导数作为指导优化的方向, 但所使用的导数阶数有所不同。假设被优化的目标函数为

梯度下降法又称最速下降法, 使用一阶导数即梯度作为优化指导, 是一种一阶迭代算法。它的变种随机梯度下降被广泛应用于深度学习模型的优化 (Rumelhart 等, 1986)。在梯度下降迭代的过程中, 每一步的方向都依赖于梯度的负方向, 这使得拟合的过程将快速前往优化函数的局部最小值:

$$\theta_{t+1} = \theta_t - \alpha g_t, \quad (2.22)$$

而

$$g_t = \nabla J(\theta_t), \quad (2.23)$$

其中, $J(\theta)$ 为待优化的目标函数, θ 为被优化的参数, α 控制了每一次迭代步长的大小。 α 与 θ_0 的取值对于梯度下降优化的成功与否至关重要。若 α 取值太小, 则收敛速度太慢; 若 α 取值太大, 可能会跳过局域最小值区域, 导致拟合失败。而若 θ_0 取值不当, 整个优化过程会囿于不正确的局域坑内, 无法找到整体最小值。

但在某些复杂的函数中, 梯度下降会因为梯度只考虑了泰勒展开的第一项而导致优化效率低。此时, 牛顿法因为支持泰勒展开第二项所携带的信息, 就变得更加适合了:

$$\theta_{t+1} = \theta_t - H_t^{-1} g_t, \quad (2.24)$$

其中 H_t 为海森矩阵。牛顿法的性能非常优秀, 但对于数据量较大的情况, 计算海森矩阵将带来极大的计算代价。因此, 高斯-牛顿法使用雅克比矩阵作为海森矩阵的近似:

$$\theta_{t+1} = \theta_t - (J_k^T J_k)^{-1} J_k e_k, \quad (2.25)$$

其中 J_k 为雅克比矩阵, e_k 为表示方向的单位向量。高斯-牛顿法的这一近似很好的解决了计算量的问题, 但无法忽视雅克比近似的海森矩阵不正定、二阶的泰勒展开的本质要求当下状态非常接近距离局域最小值等要求。针对高斯牛顿法的以上问题, LM 算法利用信赖域半径 μ , 巧妙地将梯度下降与高斯牛顿法结合在一起:

$$\theta_{t+1} = \theta_t - (J_k^T J_k + \mu I)^{-1} J_k e_k, \quad (2.26)$$

其中 I 为对角矩阵。当 μ 比较小时, 信赖域较小, 因此需要高斯-牛顿法针对小步长进行优化; 当 μ 比较大时, 认为信赖域较大, 因此大胆使用梯度下降进行大步长的空间搜索。并且, 由于海森矩阵可用于近似协方差矩阵, 从而计算拟合误差, LM 算法被广泛应用于天文界的拟合需求 (Peng 等, 2002; Erwin, 2015)。

虽说有细节上的区别, 但基于导数的方法, 其核心是一致的, 即舍精度而求速度。由于二阶导数的算法要求, 在以上列举的导数方法中, 只有梯度下降与 LM 算法是稳健的。但 LM 算法较之梯度下降速度较快的同时, 它也需要相对较大的计算代价。这虽然在串行拟合星系的时候远未触及计算机的极限, 但在使用 GPU 并行拟合星系时, 内存将先于 GPU 算力达到极限, 无法很好地利用其性

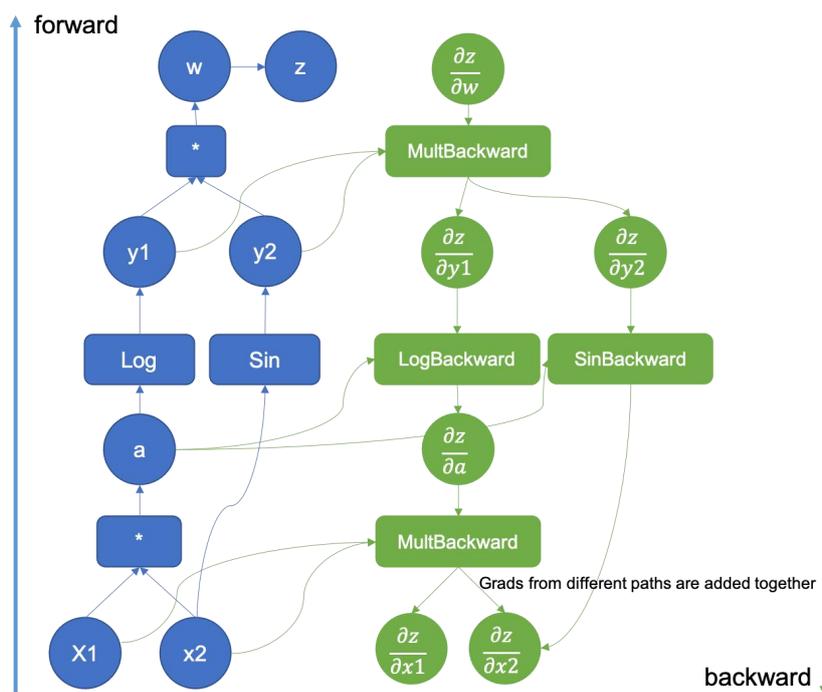


图 2.9 反向传播的动态梯度图，图源<https://pytorch.org/>。

能。因此，在考虑 GALMOSS 并行化快速拟合星系面亮度轮廓的目的，低计算开销成为我们选择梯度下降法最重要的原因。

然而，决定整体优化时间的因素除了算法本身之外，合适的起始点也十分重要。非全空间搜索的非线性优化算法，其无法逃脱寻找的是局域最小值的本质。然而，当起始点在全局最小值附近时，算法往往能寻找到正确的局域最小值——即全局最小值。这不仅仅能减少迭代的轮次，还能提升结果的正确性。但当起始点非常遥远，优化的过程大概率会囿于不正确的局域最小值。且收敛的轮次也会远超前。

2.3.3 GPU: 参数优化的加速原理

PyTorch 利用 GPU，针对梯度下降法进行了大量的优化。其中最重要的优化便是高度集成的自动反向求导的功能。对 GALMOSS 来说，只需要搭建好正向的计算过程，PyTorch 会自动进行从拟合优度指标到轮廓参数的求导，而无需进行解析的计算。这不仅减少了代码开发的繁琐程度，更是大大减少了用户自定义面亮度轮廓模型的门槛。

在 PyTorch 中，自动求导是利用动态计算图计算复合函数偏导数的链式法则实现的。链式法则可以将复合函数进行拆分，并在计算导数时进行重新组合，可以简化复合导数的导数计算。例如，当

$$z = f(y), \quad (2.27)$$

$$y = g(x), \quad (2.28)$$

则 y 关于 x 的导数可以表示为

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}. \quad (2.29)$$

动态计算图所承担的便是对复杂函数进行拆解与组合的过程。图2.9表示了一个简单函数 $f(x, y) = g(\log(x_1 \cdot x_2) \cdot \sin x_2)$ 的动态计算图的生成过程。图中，蓝色的模块展示动态的正向过程，即在函数运行的同时，计算图中便同时在计算图中添加对应的张量与函数操作。例如，当计算 $a = x_1 \cdot x_2$ 时，这一操作首先被蓝色模块记录。而由于 Torch 内部内置了基本函数的导数，因此在梯度图上，还同步计算好了 $\frac{\partial}{\partial x_1} a$ 与 $\frac{\partial}{\partial x_2} a$ 。接下来，torch 将进行 $y_1 = \log a$ 与 $y_2 = \sin x_2$ 的计算，并在绿色模块代表的反向图中添加 $\frac{\partial}{\partial a} y_1$ 与 $\frac{\partial}{\partial x_2} y_2$ 。类似的，在 $w = y_1 \cdot y_2$ 与 $z = g(w)$ 的计算过程中，torch 也会自动在反向图中添加 $\frac{\partial}{\partial y_1} w$ 、 $\frac{\partial}{\partial y_2} w$ 与 $\frac{\partial}{\partial w} z$ 。在所有前向步骤完成后，后向图将被一次性执行。torch 将按照绿色的模块，以链式法则一层层由输出节点 z 往前传播梯度，直到传播到输入节点 x_1 、 x_2 。一旦反向传播完成，动态计算图被即刻销毁。除了正向传播前所设置的叶子节点，其他节点的梯度都将被直接删除。

除了反向传播之外，优化器 (Optimizer) 也是 PyTorch 对梯度下降的重要优化功能。如章节2.3.2中所述，梯度下降作为一维的导数方法，当学习率固定的时候，只考虑到导数的一阶项。因此，梯度下降有速度较二阶导数方法较慢、容易囿于局域最小值的问题。因此，深度学习框架中所使用的优化器开始利用各种算法，削弱这一致命问题在导数方法中的影响。并且，优化器作为高度模块化的功能，将在反向传播指令后自动读取梯度，并依照所计算的学习率进行更新。

在 GALMOSS 的实操中，输入节点为轮廓拟合参数，也是我们所设置保留梯度的叶子结点； n 个星系的拟合优度之和 $\sum_{i=1}^n \chi_v$ 则为输出节点。在 PyTorch 中，轮廓拟合参数将被输入自动执行优化操作的优化器，执行反向传播后自动进行优化，并进入下一轮迭代。由于星系各自的 χ_v 仅为各自星系所对应轮廓拟合参数的函数，反向传播的过程互不影响。

2.4 参数的误差估计

然而，当使用梯度下降来进行天文领域的优化问题，如面亮度轮廓拟合，会受限于梯度下降的另一个致命缺陷：拟合误差并不天然给出。这在其他大量使用梯度下降的领域，如深度学习领域的使用场景，却是天文学家们弃用梯度下降的一个重要原因。因此，梯度下降很少作为默认算法出现在天文领域的拟合程序中。即使面亮度轮廓领域的最新工作 AstroPhot 包含了梯度下降，也依然没有给出针对梯度下降的误差估计方法。

在 GALMOSS 中，我们给出了协方差矩阵与 Bootstrap 两种误差估计的方式，补齐了使用梯度下降进行优化参数的缺口。在章3.3.2中，我们将给出两种误差方法在实际使用时的比较。



图 2.10 在自举重采样法中对星系图像进行重采样。

2.4.1 协方差矩阵

最大似然估计的背景下，海森矩阵常用于近似信息矩阵 (Pawitan, 2001)。而在基于高斯分布假设与使用负对数似然函数时 (Hogg 等, 2010)，信息矩阵的逆矩阵即协方差矩阵。由于协方差矩阵的对角线可用来计算当前最佳优化结果的拟合误差，因此，由海森矩阵出发，我们即可进行参数的优化估计。

然而，由于海森矩阵作为二阶矩阵，其在大批量拟合的前提下需要极高的计算代价。因此在 GALMOSS 中，我们选择使用雅可比矩阵来近似海森矩阵 (Gavin, 2022)。

$$\sigma_p = \sqrt{\text{diag} [J^T W J]^{-1}}, \quad (2.30)$$

其中 J 是雅可比矩阵， W 为充填了 $W_{i,j} = 1/\sigma_{i,j}^2$ 的对角矩阵。通过利用协方差矩阵估计为参数提供误差估计， σ_p 代表了参数 $1-\sigma$ 的置信水平。

由于高斯分布的假设，根据协方差矩阵来估计参数的拟合误差，其本质是传递了星系图像在观测时的统计误差。因此，当星系的面亮度轮廓分布实际上不符合我们所选模型分布时，协方差往往会低估参数真实的拟合误差。其次，协方差矩阵估计参数的拟合误差的正确性还依赖于观测误差图像的正确性。

2.4.2 自举重采样

自举重采样法是由 Efron 于 1992 年提出 (Efron, 1992)。它采用了经典频率派的思想，使用有放回的随机采样实现多次的测量分析，来预估如方差、均值等属性。自举重采样法与使用协方差矩阵不同，它的基本思想是经典统计派的思想：

如图 2.10 所示，在 GALMOSS 中，我们将对不包含在掩膜内的像素进行有放回的多次采样（默认值为 $M = 100$ 次）。随后，我们将从图像拟合阶段结束后得到的最优参数作为起始点，对 M 次采样得到的星系图像进行重拟合。由于面亮度轮廓拟合的过程对星系的中心位置非常敏感，因此若重采样过程并未采样到星系中心的像素会导致拟合过程的波动非常大。为避免重拟合的过程出现失败进而影响总体拟合误差的估计，我们将星系中心固定，不作为可优化参数参与优化过程。

在重拟合结束后，我们可以根据 M 次拟合的参数分布进行误差估计：

$$\sigma^2 = \frac{1}{M} \sum_{i=1}^M [m_i - \bar{m}]^2, \quad (2.31)$$

其中 \bar{m} 是估计参数的平均值。

自举重采样法的优点在于简单与稳健。通过自举重采样法估计的参数误差能够很好地将观测误差、模型选取的误差都考虑在内。然而，由于其正确性依赖于一定的重采样次数（样本量需要足够大），其计算开销也是不可忽视的缺陷。

第3章 性能评估：面亮度轮廓重拟合实验

为了评估 GALMOSS 的性能，我们将在本章中使用 GALMOSS 对 Sloan 数字巡天的星系图像进行拟合，进而评估其准确性与速度。在进行准确性的评估时，我们使用目前权威的星系面亮度拟合程序 GALFIT 所拟合的星系测光星表结果作为参照组。但因为测光星表并不释放拟合的时间，我们需要重新测算 GALFIT 拟合相同数据量的星系图像所需要的时间。因此我们将在使用相同的数据集的前提下，使用 GALFIT 再次拟合所选星系样本，从而实现速度上的评估。

3.1 数据

3.1.1 SDSS 的测光图像

Sloan 数字巡天旨在为宇宙绘制详细的天体“地图”。它利用位于美国新墨西哥州阿帕奇角天文台的 2.5 米光学望远镜进行测光与光谱观测。2009 年后，该望远镜测光任务结束，完整的成像数据发布覆盖了超过 35% 的天空 (Ahn 等, 2012)，着重覆盖北银冠与南银冠地区，包括了约 930,000 个星系、120,000 个类星体和 225,000 个恒星。Sloan 数字巡天的测光观测包括从紫外到近红外的五个波段，分别为 u, g, r, i, z ，其对应中心波长为 3550Å, 4770Å, 6230Å, 7630Å, 9130Å。当视宁度为 1 角秒时，所能达到的 5σ 探测极限分别为 22.3, 23.3, 23.1, 22.3 与 22.8。Sloan 数字巡天的成像相机的像素大小约为 $24\mu m$ ，即 0.396 角秒。

在 SDSS 中，CCD 的大小为 2048×2048 像素，所接收的光子数均校准到以 nanomaggies 为单位的流量 f ，从而使得流量的数值范围和精度更适合 SDSS 所观测的天体亮度范围。一个以 nMgy 为单位的星系与经典的 AB 星等的近似换算公式如下：

$$m = 22.5 - 2.5 \log_{10} f, \quad (3.1)$$

即 SDSS 的星等零点为 22.5。而 f 经过标准源映射，其单位与一般的央斯基单位 (Jy) 有如下关系：

$$f \sim 3.631 \times 10^{-6} Jy. \quad (3.2)$$

SDSS 的测光数据中并不直接包含有当前曝光的观测误差以及点扩散函数，但其在观测文件¹中提供了天空 (counts)、校准图像 (nanomaggies count⁻¹)、增益、暗方差以及半高全宽²，可以根据利用泊松分布的方差以及误差传递公式计算观测误差³，并利用高斯分布对点扩散函数进行模拟。

¹<https://dr15.sdss.org/sas/dr15/eboss/photoObj/>

²<https://dr15.sdss.org/sas/dr15/eboss/photoObj/301/>

³https://data.sdss.org/datamodel/files/BOSS_PHOTOOBJ/frames/RERUN/RUN/CAMCOL/frame.html

3.1.2 MANGA 的测光与形态星表

在结束测光观测后，为准确的测量星系和类星体的大尺度分布，SDSS 选择了两个无偏星系样本进行后续的光谱观测。一个是根据流量进行筛选的主星系样本（消光改正后 r 波段的 Petrosian 星等 $r < 17.77$ ），另一个是根据颜色与流量选择的亮红星系样本，其 r 波段的 Petrosian 星等延伸至 19.5。随后，SDSS 旨在观测近邻星系的子项目 MANGA 从主星系样本中进一步选择了 10,293 个星系作为其样本。它们是根据恒星质量和半光半径挑选而出，比恒星质量 $\log M_*/M_\odot \geq 9$ ，且具有相对均匀的恒星质量分布 (Collaboration, 2022)。

- MaNGA PyMorph 测光增值星表 (MPP-VAC)

MPP-VACfischer2019sdss 提供了上述样本在 g, r, z 三个波段的测光参数，包括星系方位角、轴比、星等、半光半径以及描述面亮度轮廓变化的 sersic 指数。MPP-VAC 使用 SDSS 第 7 次数据释放的测光数据拟合了两种面亮度轮廓模型，分别是单成分模型和双成分模型。在单成分模型中，星系的面亮度轮廓由 sersic 模型完全描述。在双成分模型中，星系的面亮度轮廓由两个 sersic 轮廓进行分解。其中一个 sersic 轮廓的 index 指数被设定为 $n = 1$ ，意在表示星系中的盘成分；另一个 sersic 轮廓的 index 指数在 $0 \sim 1$ 的范围内自由拟合，意在表示核球成分。

正如其名，MPP-VAC 星表使用了基于 Python 的拟合程序 PyMorph 来拟合星系的测光参数。PyMorph 作为一个脚本程序，其作用是粘合 SExtractor 与 GALFIT。其中，SExtractor 的作用是提供初始参数与掩膜图像。在指定一定的超参后，SExtractor 将对输入的图像源的检测、测量并自动输出文档。其测量的结果包括背景估计、光度测量、形态参数测量等等，可以直接作为合适的初始参数，提高 GALFIT 拟合成功的概率。除了初始参数之外，SExtractor 还可以输出掩膜图像，对背景、源的区域进行标注。随后，PyMorph 将 SExtractor 输出的结果合并其他图像数据集（星系流量图像、星系流量误差图像、星系点扩散函数图像）一并输入 GALFIT，并由 GALFIT 得到最终的拟合结果。在 PyMorph 中，单成分拟合直接使用了 SExtractor 输出的各项初始参数；在双成分拟合的过程中，核球与盘共享星系中心、轴比、半光半径、星等等参数的初始值。但在 sersic 指数初始值的选择上，PyMorph 默认设置 4 为拟合起点。

鉴于 GALFIT 在星系面亮度轮廓研究中的广泛使用，我们选取其单成分模型的结果作为对照组，与 GALMOSS 的结果进行对比。

- MaNGA 深度学习形态增值星表 (MDLM-VAC)

MPP-VAC 星表中包含的星系样本，其形态分类可以通过 MDLM-VAC (Fischer 等, 2019) 来获取。该形态分类并非由人眼识别，而是通过 Domínguez Sánchez 等 (2018) 训练的神经网络得到。该神经网络由 JPG 格式的 SDSS 第七次数据释放的 RGB 图像，NA10 数据集 (Nair 等, 2010) 提供的星系形态参数 T ，以及星系动物园第二代分类星表的结果进行训练 (Willett 等, 2013)。当输入并不参与神经网络训练过程的图像时，该网络可以输出 T type 值，分类为 S0 星系、侧向星系、含棒星系、并合星系、含核球星系与雪茄星系的概率，以及额外添加的两类标签：

Visual Class (VC) 与 Visual Flag (VF)。前者用于区分椭圆、S0 与旋涡星系，即当 $VC == 1$ 时为椭圆星系， $VC == 2$ 时为 S0 星系，当 $VC == 3$ 为旋涡星系。后者用于区分 VC 的值是否可靠， $VF == 0$ 代表 VC 得到的分类是较为确定的，而反之则不确定。

根据 MDLM-VAC 所提供的分类建议，我们以表 3.1 展示的分类方法对星系的形态分类的参考。具体将应用于下一小节 Sersic 指数偏移的分析。

表 3.1 星系分类方法
Table 3.1 Galaxy classification scheme

Type	Method
LTG	$T_type > 0$
ETG	$T_type < 0$
E	$T_type < 0, P_{S0} < 0.5, VC == 1, VF == 0$
S0	$T_type < 0, P_{S0} > 0.5, VC == 2, VF == 0$

3.1.3 数据集的制作

与其他传统方法类似，GALMOSS 的运行同样需要图像数据集与初始参数。为了完整地将星系囊括在图像中，我们将星系图像数据集的大小规定为 128×128 。因此在选择样本的过程中，我们去除了位于 SDSS 的 CCD 边缘的星系，因为这些星系无法完整地裁剪出大小为 128×128 的图像。其次，我们还去除了 MPP-VAC 星表中单成分模型拟合失败的星系（在星表的对应位置显示为 999 或 NaN）。最终，经过筛选后的最终星系样本包含 8,289 个星系。

为了确保整个样本具有良好的信噪比，我们选择了 SDSS 的 g 波段图像作为我们的实验目标。随后，我们根据波段与星系样本的需求下载对应 SDSS 的原图像文件，并依照 3.1 所介绍的流程进行图像的裁剪、观测误差图像与点扩散函数图像的制作。在点扩散函数图像的制作的过程中，我们将图像的大小确定为 40×40 。参考 MPP-VAC 的拟合过程，我们也使用了相同的流程来生成掩膜图像和拟合的初始参数。但与其选择不同的是，我们使用 SExtractor 官方文档推荐的 Sersic 指数近似作为其初始参数，即半光半径 R_e 与 Kron 半径的比值。图 3.1 展示了一个星系对应图像数据集的示例。

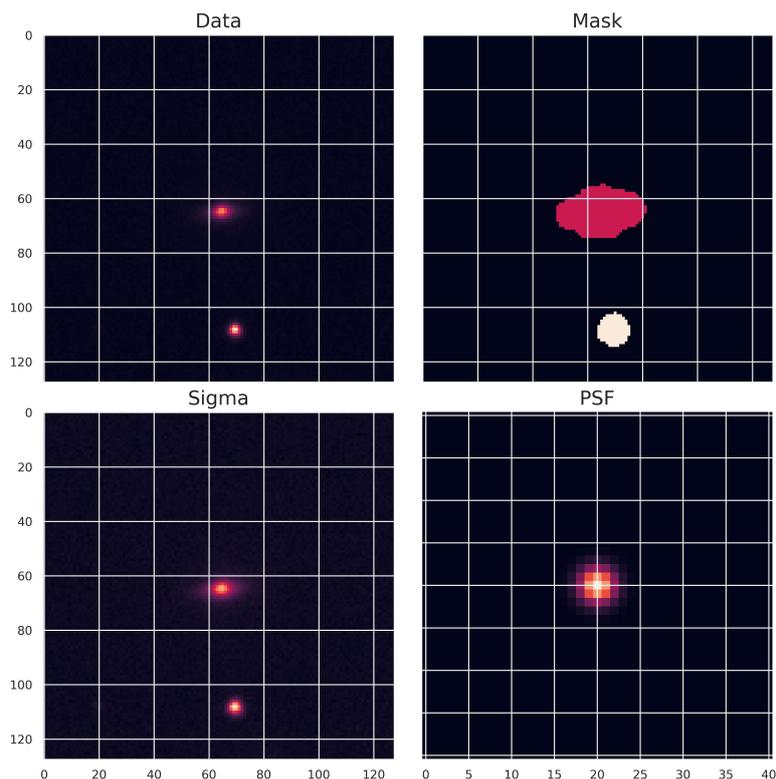


图 3.1 星系图像数据集的实例。

3.2 实验设置

3.2.1 GALMOSS 重拟合实验设置

Shen 等 (2022) 在对 SDSS 同样由 GALFIT 拟合得到的不同星表进行对比时发现, 当进行核球、盘成分拟合的时候, 不同的初始条件、拟合限制条件、PSF 平均值的选取、拟合的策略不同, 会导致结果出现较大的差异。然而, 在单 Sersic 拟合的条件下, 不同星表的结果没有太大的差别。因此, 为了尽量避免引入其他误差项, 使得结果的差异只来源于不同软件的内禀差异, 我们选择使用单成分拟合 (Sersic 轮廓) 来重拟合所选样本。为了还原拟合参数限制, 我们不加入用于限制 Boxy/Disky 的参数 B , 并同样给 Sersic 指数的拟合范围加上 $0 \sim 8$ 的硬边界限制。同时, 为了更好的拟合效果 (Robotham 等, 2017), 我们使 Sersic 指数在 \log 参数空间中进行拟合。

为了确定拟合的最大迭代数, 我们随机选择了 1,000 个星系进行测试。测试星系组的残差回归图表示, 在迭代数大于 800 时, 残差基本收敛并进入稳定状态 (图3.2)。为了在拟合速度与拟合准确性之间取得平衡, 我们最终选择 $i = 1000$ 为最大迭代数。在批数量 Batch size 的选择上, 我们选择了能最大化利用 GPU 算力, 且为 2 的倍数的 1024。在实验中, 我们发现当批数量大于 1024 时, GPU 已经达到满算力。在满算力情况下, 尽管 GPU 运行内存仍支持更大的批数量, 但实际使用时 GPU 会在内部将批数量进一步分开, 成为串行操作。在这种情况下, 大批数量对速度的提升没有帮助。而如若批数量为 2 的倍数, 能够提升 GPU 的

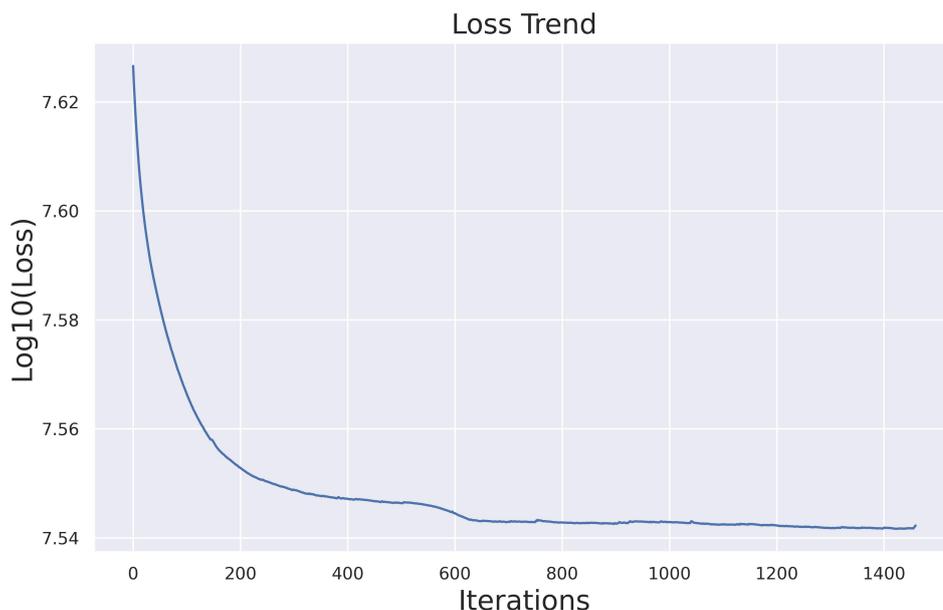


图 3.2 测试星系的 loss 曲线。

使用性能。在优化器的选择上，我们选择了 Adam。其作为深度学习运用最为广泛的优化器之一 (Kingma 等, 2020)，其性能足够稳定，并能帮助梯度下降算法一定程度上逃脱囿于错误的局域最小值的问题。

在实验过程中，我们将记录最终的最佳拟合值以及模型的拟合时间。除了拟合全样本星系的总体时间之外，我们还进行了拟合单星系、批数量为 100、批数量为 1000 的实验并记录时间。最佳拟合值将用于拟合准确性的评估，而拟合时间将用于后续速度性能的评估。但由于 GALFIT 拟合时间数据的确实，我们利用 GALMOSS 拟合所使用的数据集，运行脚本⁴进行拟合。为了客观进行速度比较，我们将线程数定为 1，使 GALFIT 进行串行优化。

实验的具体环境为 python3.9 和 torch2.0.1+cu117；所用硬件为 CPU 为 2.2GHz Intel Xeon Silver 4210；显卡为 NVIDIA A100 (80 GB)；系统为 Ubuntu Linux 18.04 64 bits；RAM 为 3.9TB。

3.2.2 统计指标

为了评估 GALMOSS 结果与 GALFIT 结果的一致性，我们引入 R^2 决定系数进行量化。

R^2 决定系数

R^2 决定系数由 Wright (1921) 首次发表，常用于回归分析的评估。其定义如下：

⁴<https://github.com/Grillard/GalfitPyWrap>

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}, \quad (3.3)$$

其中 y_i 为 GALFIT 的拟合结果, \hat{y} 为 GALMOSS 的拟合结果, \bar{y} 为 GALFIT 拟合结果的均值。在上式中, 分子为残差的平方和, 分母为总平方和 (与数据的方差成正比)。在拟合值与对照值完全匹配时, 有决定系数 $R^2 = 1$ 。在一个拟合值等于对照值平均值的基准情况下, 有决定系数 $R^2 = 0$ 。在一些更差的拟合情况下, 将会有 $R^2 < 0$ 的极端情况。在回归分析的评估中, 决定系数 R^2 能比 MSE 等常用的评价指标更适合在不同数量级的参数之间进行直观的对比。因为后者的值域是任意的, 其具体的大小与参数的数量级有关; 而决定系数的值域基本在 $0 \sim 1$ 之间 (除了极端的拟合失败情况)。

归一化的中值绝对偏差

中值绝对偏差 (median absolute deviation, MAD) 用于衡量数据的离散程度, 其概念在 Gauss (1816) 关于确定数值观测精度的工作中被首次提及。在定义绝对偏差后:

$$\Delta y = \frac{y_i - \hat{y}_i}{y}, \quad (3.4)$$

我们可以得到 MAD 的定义如下:

$$\text{MAD} = \text{median}(|\Delta y - \text{median}(\Delta y)|). \quad (3.5)$$

NMAD (normalized median absolute deviation) 为归一化后的 MAD, 使得在正态分布的情况下, NMAD 的数值等于标准差:

$$\text{NMAD} = 1.4826 \times \text{MAD}. \quad (3.6)$$

NMAD 相较于标准差更加稳健, 受少数异常值的偏差影响较小, 因此它也更适用于无均值或方差的分布, 例如柯西分布。

异常值比例

在本文中, 异常值比例定义为 $\Delta y > 15\%$ 的比例。这一项指标考察了数据集中严重偏移正常范围的异常值, 可能是由拟合失败、星系不能由所选模型很好表示等原因产生。

3.3 拟合结果与讨论

在上述实验条件的设置下, GALMOSS 耗费了 10 分钟的时间完成全样本的面亮度轮廓拟合, 并将每个星系的图像输出与最佳拟合值输出在给定路径下。在下文中, 我们将从拟合准确性、拟合的速度两方面进行 GALMOSS 性能的评估。

3.3.1 拟合准确性与讨论

我们读取最佳拟合值所保存的 FITS 文件，并提取星系样本各自的拟合优度 (χ_μ) 展示在了图3.4中。如图所示，大部分星系都处在 $\chi_\mu < 2$ 的区间中，表明其拟合的结果比较成功地还原了星系的面亮度轮廓分布。除此之外，还有部分星系的 χ_μ 值偏大，说明单成分拟合无法成功还原其亮度分布。因此，我们进一步挑选了 $\chi_\mu < 1.01$ 与 $\chi_\mu > 7$ 的星系各三个，展示在图3.3中。其中左边的三列分别为 $\chi_\mu < 1.01$ 的星系的流量图片、模型图片与残差图片；右边为 $\chi_\mu > 7$ 的星系的流量图片、模型图片与残差图片。可以看出，拟合较为成功的星系，成分比较单一，基本可以用一个 Sersic 轮廓来描述。其残差也呈现高斯分布的误差起伏。而拟合结果无法很好形容星系流量分布的样本，他们的成分比较复杂。如第一行的星系，可以明显看出是一个侧向星系，由盘与核球两个成分组成；第二行的星系则是一个并合星系，其形态并不规则；第三行的星系是一个拥有絮状旋臂的旋涡星系。从残差的结果可以看到，被过度剪除和没有被拟合上的流量占据主导，高斯分布的背景无法显现在残差图像中。因此，Sersic 轮廓作为一个平滑的轮廓模型，其表达能力非常有限，无法很好地对不规则结构、絮状旋臂进行描述。在明显拥有双成分结构的星系中，单 Sersic 模型无法兼顾核球与盘的面亮度变化表现。

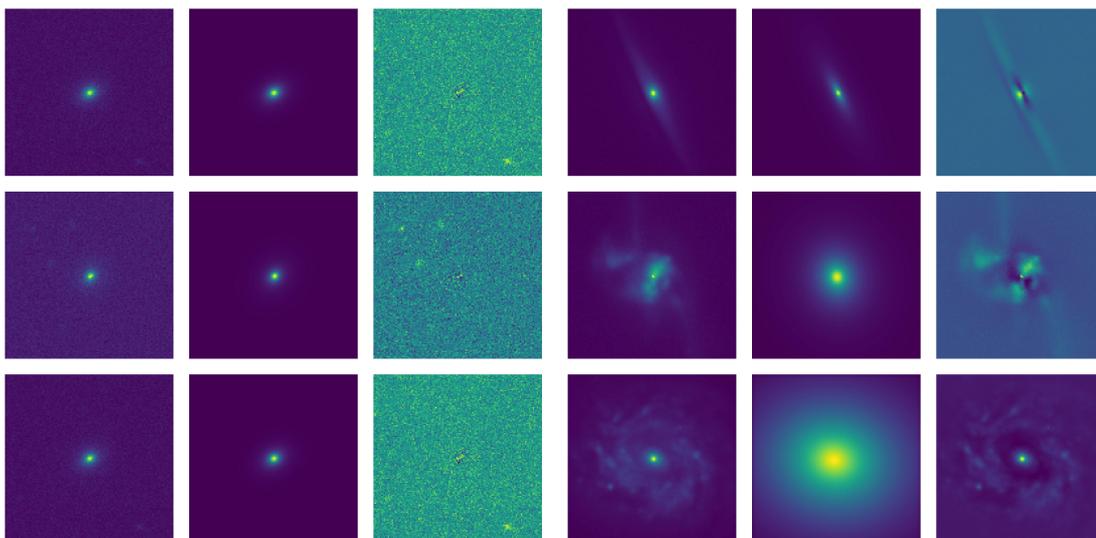


图 3.3 GALMOSS 拟合结果的视觉检查。

图 3.5 显示了 GALFIT 拟合结果 (MPP-VAC 星表) 与 GALMOSS 拟合结果的对比，包括星等 (m)、方位角 (θ)、轴比 (q)、有效半径 (r_e) 和 Sérsic 指数 (n)。散点的位置由 GALFIT、GALMOSS 拟合值决定，散点的颜色由其所在位置散点的密集程度所决定。图上所显示的 R^2 决定系数为该参数拟合结果一致性的量化。此外，左边和上方的边缘分别显示了垂直和水平轴上测量值的分布直方图。这些直方图展示了两种方法各自测量结果的集中趋势和分布范围。全部的统计指标

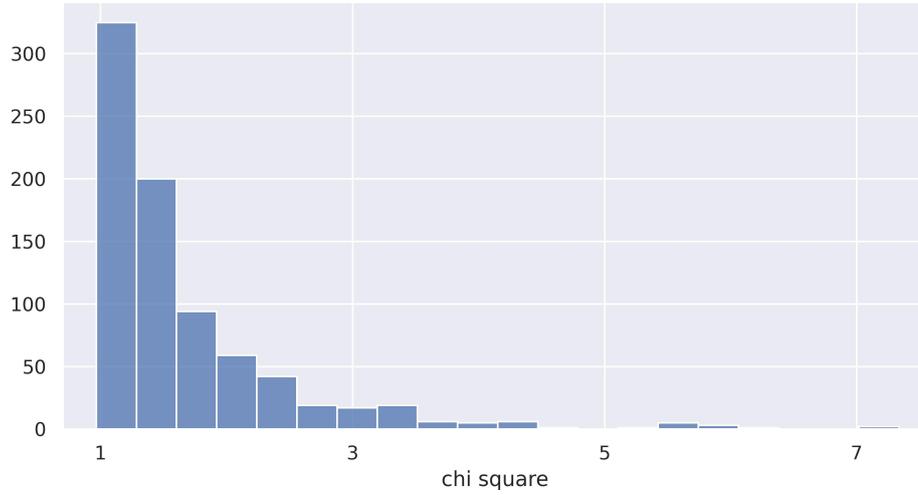


图 3.4 对 8,329 个星系进行拟合后的 χ^2 直方图。

(包括 R^2 决定系数, NMAD 与异常值比例) 被展示在表 3.2 中。

表 3.2 GALMOSS 拟合结果的评估

Test	mag	q	n	θ	r_e
R^2	0.9485	0.9628	0.7943	0.7943	0.7663
异常值比例	0.0036	0.0304	0.1278	0.0289	0.0755
NMAD	0.0011	0.0177	0.2472	0.0845	0.1183

作为星系几何参数的星等、方位角和轴比的结果显示出很强的线性相关性, 表明 GALMOSS 的结果与 GALFIT 的结果高度一致。星等在两个程序的拟合结果弥散非常小, R^2 约为 0.95, 异常值比例约为 0.004, NMAD 约为 0.001。两种结果的分布也较为一致, 均大约在 14 等到 18 等之间, 在主要集中在略高于 16 等的位置, 在 15 等左右也都有一个小的聚集。轴比的结果也非常一致, R^2 约为 0.96, 异常值比例约为 0.0043, NMAD 约为 0.02。轴比的分布较为均匀, 约等于 0.8 的星系占据大多数。在轴比较大的星系中, 其弥散也稍大一些。比较特殊的是方位角的拟合。从图中可以看出, 颜色较深 (密集程度非常高) 的散点与参考线的弥散是最小的, 但在其他区域还零散分布着非常多的散点, 即一致性较差的点, 因此, R^2 受到了一定的影响, 约为 0.79。但在其他统计指标上, 方位角表现优秀, 其异常值比例约为 0.03, NMAD 约为 0.08。其中, 在 $\theta_{GALMOSS} \sim 90$, $\theta_{GALFIT} \sim -90$ 与 $\theta_{GALMOSS} \sim -90$, $\theta_{GALFIT} \sim 90$ 的两个端点也有部分散点聚集。一致性较差的点, 即明显偏离参考线的散点为形态较接近圆形的星系, 即 ($q \sim 1$)。在这种情况下, 星系的面亮度轮廓分布对位置角 (θ) 的变化相对不敏感。而在端点处的散点, 是由于在面亮度轮廓拟合的时候, θ 的值在 -90 与 90 两种情况下是等价的。因此, 方位角的拟合依然是成功的。

星系的面亮度轮廓参数, 即半光半径 r_e 和 Sersic 指数 n 是相对来说较难精

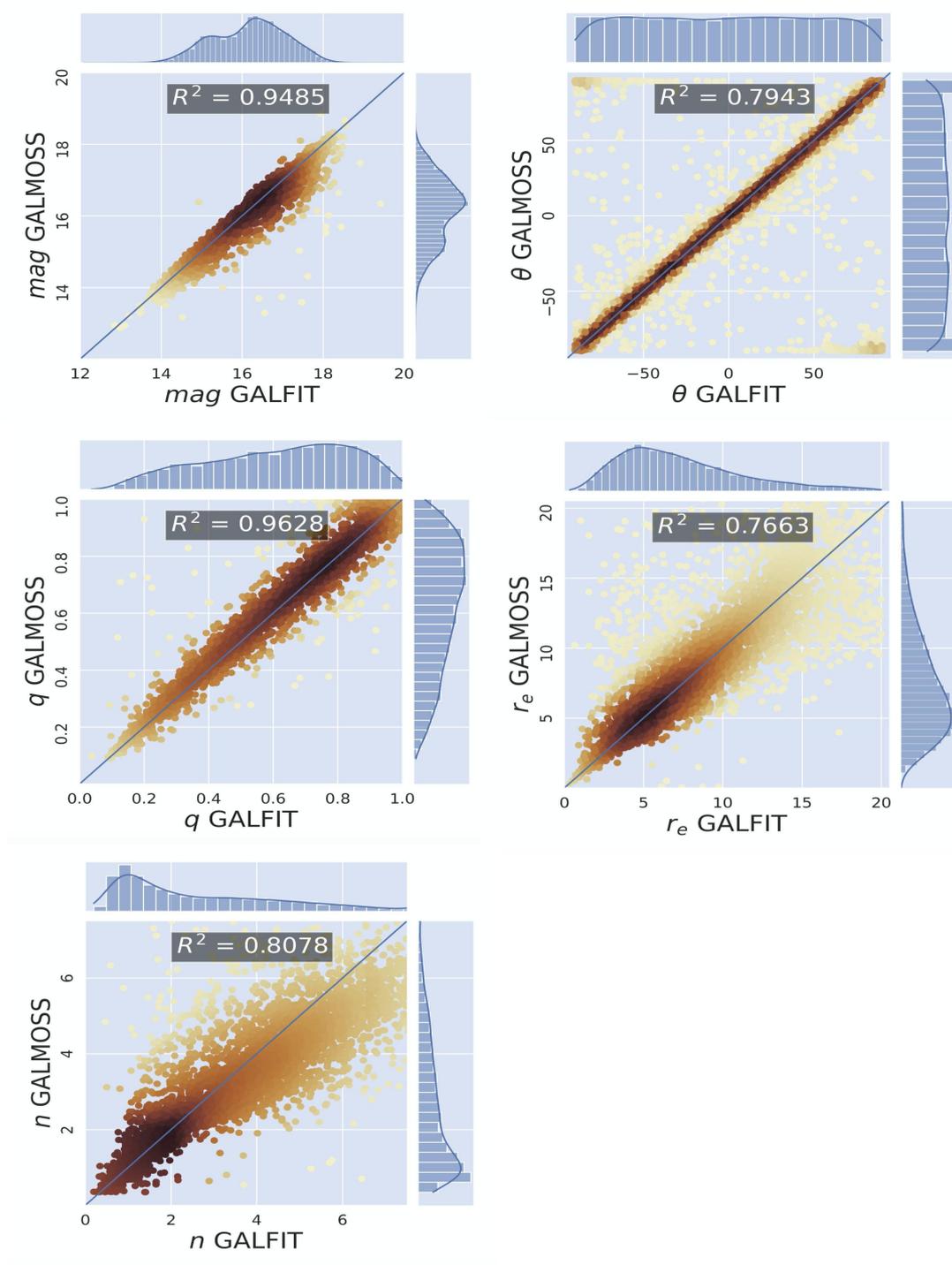


图 3.5 此图比较了所选 ~8,000 个 SDSS 星系的测量 Sérsic 剖面参数，以及其与 GALFIT 结果 (MPP-VAC 星表) 比较的 R^2 决定系数。子图分别显示了星等 (m)、方位角 (θ)、轴比 (q)、有效半径 (r_e) 和 Sérsic 指数 (n) 的拟合结果。点的颜色说明了它们分布的密度。蓝线显示 galfit 结果完美符合 galmoss 结果的参考线。顶部和右侧的直方图分别说明了 galmoss 和 galfit 测量值的分布。此图显示，GALMOSS 拟合的所有参数都与 GALFIT 拟合的结果具有高度的一致性。

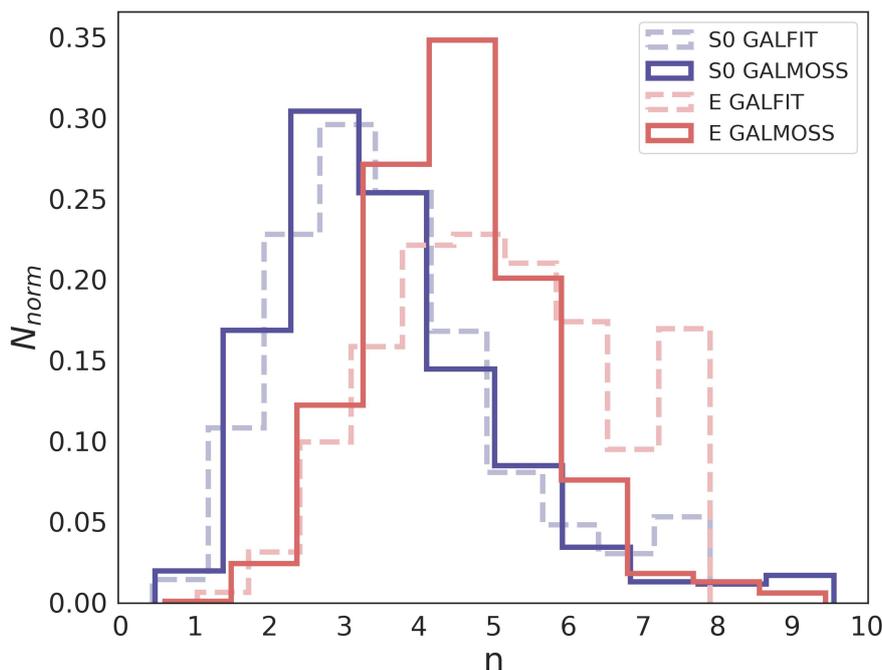


图 3.6 S0 星系、椭圆星系不同来源拟合值的直方图。紫色代表了 S0 星系 Sersic 指数 n 的分布，红色代表了椭圆星系 Sersic 指数 n 的分布；而虚线则为 GALFIT 拟合的结果，实线为 GALMOSS 拟合的结果。由图可得，椭圆星系对拟合 n 值的不一致贡献最大。

确拟合的 (Trujillo 等, 2001)。从 R^2 的数量级上也可以看出，面亮度轮廓参数对应的 0.77 与 0.81，相对几何参数的数量级相对较低。同时，异常值比例与 NMAD 也有类似的表现： r_e 和 n 的异常值比例约为 0.08 和 0.13，NMAD 约为 0.12 和 0.25。但在上图中，它们依然表现出较好的线性关系。半光半径 r_e 的分布主要集中在 5 角秒左右，弥散随着半光半径的增长而有所增加。Sersic 指数 n 是五个参数拟合结果中相对弥散最大的参数，其分布主要集中在 $n < 2$ 的位置。值得注意的是，在 Sersic 指数较大的地方 $n > 6$ ，其弥散并不对称，而是出现了一定程度的偏移。在这块区域，GALMOSS 所拟合得到的结果通常低于 GALFIT 拟合得到的结果。这种差异可能是图像质量的变化以及 PyTorch 中使用的优化器的算法特征导致的。接下来，我们将 $n_{GALFIT} > 6$ 与 $n_{GALMOSS} < 6$ 的区域定义为偏移区域，并对这种偏移现象产生的原因进行进一步的分析。

首先，我们确定了处于偏移区域中的星系所属的类型。通过 MDLM-VAC 星表与上文所描述的星系分类方法，我们将被样本所包含的 8,289 个星系识别为 2,849 个早型星系和 5,440 个晚型星系。我们发现，尽管早型星系仅占整个星系样本的 34.37%，但在偏差区域中占 88.78%。这证明，在偏移区域中主要起作用的应为面亮度轮廓较为平滑且紧凑的早型星系。为了进一步确定对偏移区域贡献最大的星系类别，我们将 ETG 识别为 1,891 个椭圆星系、793 个 S0 星系和 165 个未定义星系。排除未定义星系，我们提取了样本中属于椭圆星系或 S0 的星系的 Sersic 指数 n 的拟合值，并将 GALMOSS 的拟合结果与 GALFIT 的结果分别展示在直方图中。如图 3.6 所示，紫色代表了 S0 星系 Sersic 指数 n 的分布，红色

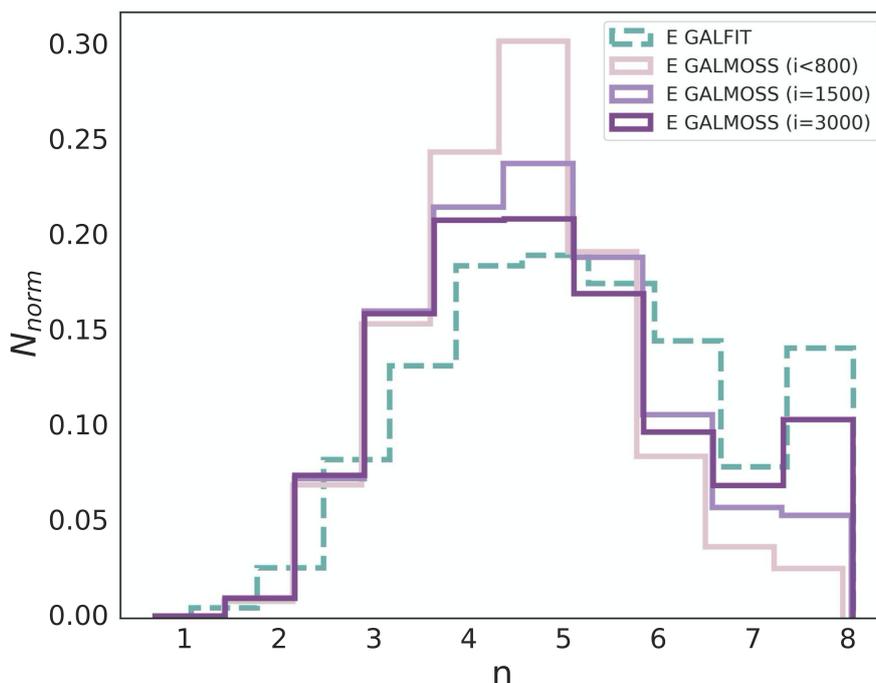


图 3.7 对椭圆星系进行不同迭代数拟合的结果。随着迭代数的增加，GALMOSS 的结果逐渐向 GALFIT 逼近。

代表了椭圆星系 Sersic 指数 n 的分布；而虚线则为 GALFIT 拟合的结果，实线为 GALMOSS 拟合的结果。从 S0 星系与椭圆星系实线、虚线的走势可以看出，S0 星系的 GALFIT、GALMOSS 结果差距并不大，没有体现出明显的系统性偏差；然而椭圆星系差异较为明显。GALMOSS 拟合得到的椭圆星系 n 值的峰值主要集中在 $n = 4 \sim 5$ 之间，在 n 较大的区间分布很少；而 GALFIT 的椭圆星系 n 分布比较平均，且在 $n \sim 8$ 时有一个异常的凸起。从上述分析我们可以得出结论，椭圆星系对拟合 n 值的不一致贡献最大。而当我们把最大迭代数设置为小于 800, 1500 与 3000 并对椭圆星系进行重拟合时，如图 3.7 所示，随着迭代数的增大，结果也逐渐向 GALFIT 逼近。由此得知，在对 n 偏大椭圆星系进行拟合时，GALMOSS 的收敛会减缓。

星系的 Sersic 指数 n 在硬性限制的区间上大量累积是大样本拟合的普遍现象 ((Blanton, M. <http://nsatlas.org/data>; Meert 等, 2015; Simard 等, 2011)，是多种不同的因素共同参与贡献而导致的。首先，观测图像的限制性是其中一个重要的原因。大多数椭圆星系，其中央的流量非常致密，变化的速率非常快；而到了星系外围，流量的变化速率快速降低，形成一个平坦的外翼。在星系中心，像素与像素之间的较大流量差异本应很好地约束 Sersic 轮廓的指数因子，但因为天文观测所使用的 CCD 的像素大小有限，导致快速变化的区域往往被压缩在非常小的像素区域内；而投影在同一个像素面积内的流量梯度变化会被直接抹平，无法对 Sersic 指数因子的优化提供很好的指导。在椭圆星系的外围，尽管所占据的像素区域非常大，但其偏小的流量导致信噪比相对较差、而且，平坦的星系外翼流量

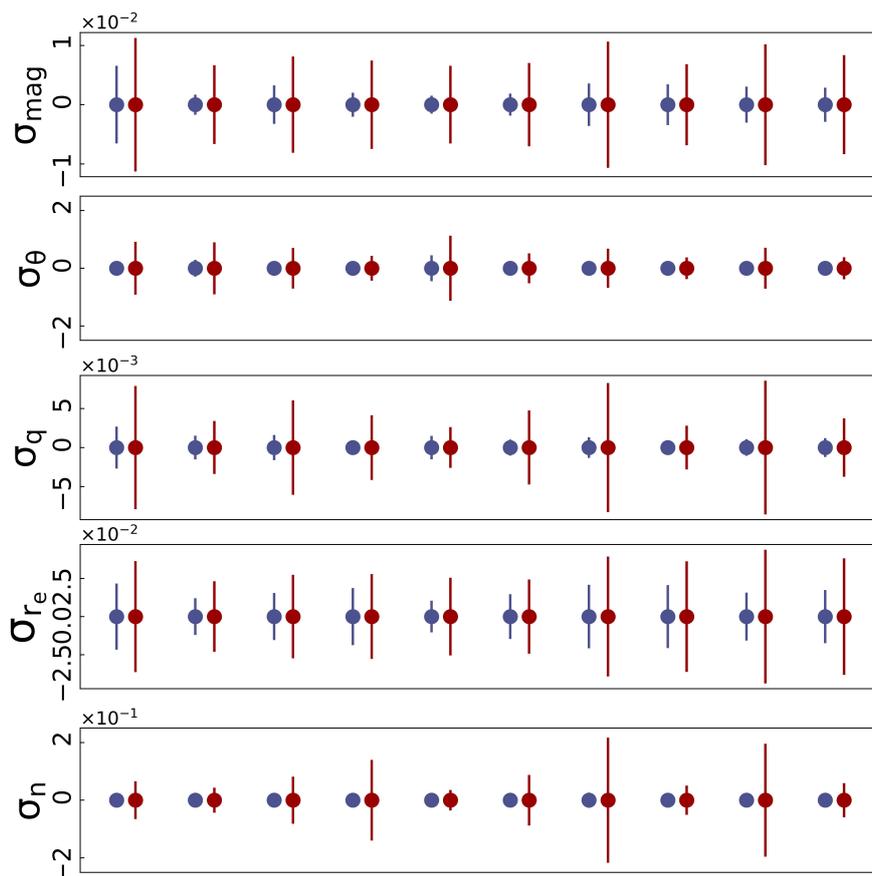


图 3.8 使用样本中随机选择的 10 个星系对 GALMOSS 内置的两种拟合不确定性估计方法进行比较：协方差矩阵（用蓝色误差棒表示）和自举法（用红色误差条表示）。通常，使用自举重采样法得出的不确定性大于使用协方差矩阵获得的不确定性。

梯度非常小， χ^2 对 n 的变化不敏感。而且，从 Sersic 公式本身出发，其不同指数因子之间，面亮度模型轮廓的区别也有所不同。当 $n \sim \infty$ ，Sersic 轮廓模型会无限趋近于斜率为 5 的幂律模型，这也使得 χ^2 无法很好地优化 n 。

另一个可能的原因则是 PyTorch 框架所支持优化器的内部算法对学习步长的削减。由 2.3.3 我们可知，优化器会使用各种算法，动态的改变和优化学习率。比如我们在实验中使用的 Adam 优化器，它支持自适应学习率、动量与 RMSprop，通过改变学习率来优化每次学习的步长。其中的自适应学习率在大梯度条件下将增拥有较大的学习率，而低梯度条件下将降低学习率以加快收敛。因此，与 GALFIT 相比， n 的灵敏度在值较大的区域会得到进一步的降低，这增加了本例中两个软件之间结果的差异。

3.3.2 拟合误差的评估

为了更好地比较 GALMOSS 所提供的两种拟合不确定性方法，我们随机选择了 10 个星系，在图 3.8 中展示了其两种不确定性的值。由于在此次评估中拟合值本身并非关注的对象，因此我们将不同星系的不同参数值统一到了 0。蓝色的点与误差棒显示的为统一后的拟合值与协方差矩阵所得到的拟合不确定性；红

色的点与误差棒显示的为统一后的拟合值与自举重采样法所得到的拟合不确定性。如图所示，自举重采样法的误差普遍大于协方差矩阵所得到的误差，表明其估计更为保守，相较于协方差矩阵能够考虑更多种类的误差。这与章2.4所述一致。

3.3.3 拟合速度的评估

在全样本的总体拟合实验中，GALMOSS 拟合 8,289 个星系大约需要 10 分钟。因为此时 GPU 的利用率已经到达上限，这可以视为在当前实验环境与实验设置下，GALMOSS 所能达到的速度极限。而与此同时，我们使用 GALFIT，利用 CPU 串行拟合相同数量的星系则需要大约 60 分钟。换言之，GALMOSS 将 GALFIT 的速度提升了大约六倍，这在大样本的前提下已经比较可观。

为了更直观地展现并行化拟合对 GALMOSS 速度提升的效益，我们将 GALFIT 的平均拟合时间作为基线，与拟合单个星系、10 个星系、100 个星系与 1000 个星系四个不同的情况进行拟合速度上的对比。可以发现，当同时并行的星系较少的时候，GALMOSS 无法体现出使用 GPU 的优势，速度较 GALFIT 较慢。然而，GALMOSS 的速度优势随着批数量大小的增加变得更加明显，直至达到 GPU 计算能力的极限。从图3.9可以看出，在批数量大于 10 个左右时，GALMOSS 的拟合速度开始超越 GALFIT，当批数量等于 1000 时，拟合的加速倍数已经逼近 6 倍。

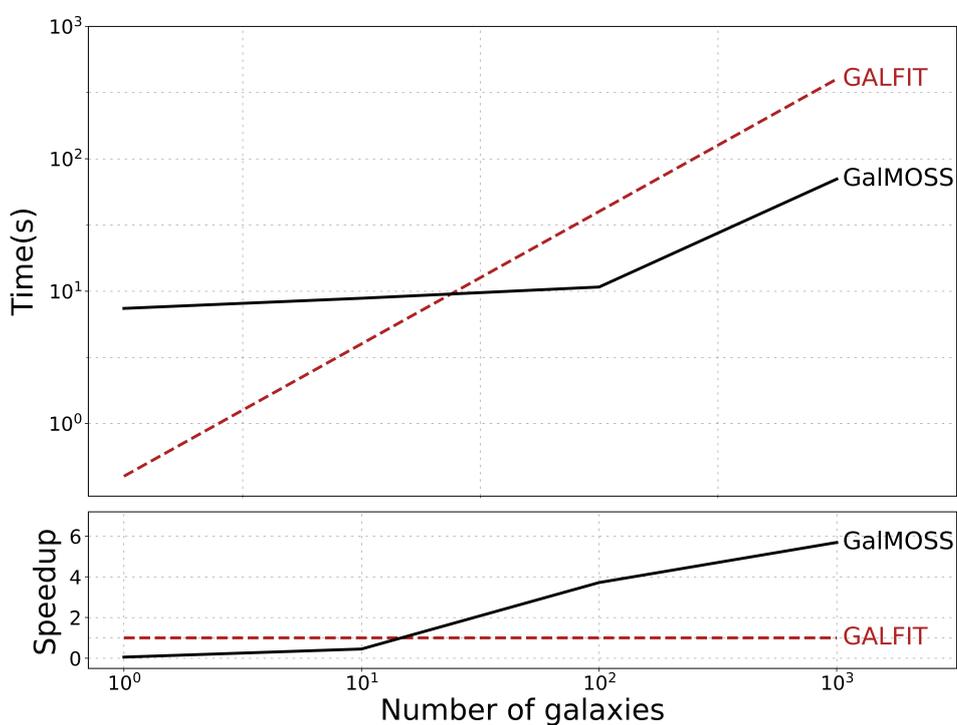


图 3.9 GALMOSS 和 GALFIT 的运行时间比较，以 GALFIT 运行时间为标准进行归一化。结果表明，当批数量稍大于 10 的时候，GALMOSS 开始超越 GALFIT 的速度。当批数量为 1000 的时候，GALMOSS 将 GALFIT 的速度提升了大约六倍。

第4章 总结、讨论与未来工作展望

4.1 总结

在本文中，我们介绍了我们的面亮度拟合软件 GALMOSS。GALMOSS 是基于 Python 与 PyTorch 框架所构建的开源软件包，专门为在大样本数据集上拟合星系面亮度轮廓而设计，非常适合即将到来的大数据时代。GALMOSS 结合了传统方法与现代方法，具有多项优势与特点。

在图像的生成阶段，GALMOSS 能以多维矩阵计算的形式，从参数批量生成二维的面亮度轮廓模型图像。在这一阶段，参数以轮廓模型公式的形式参与了计算，使得拟合过程保留了物理意义。与此同时，GPU 将对图像生成过程进行加速，大幅度缩短该阶段所需要的时间。作为面向对象的程序，尽管 GALMOSS 内置的轮廓模型只有七种，但在此阶段加入用户自定义的模型非常方便。并且，只需构建好正向的生成公式，GALMOSS 将随后自动执行图像的拟合过程。

在图像的拟合阶段，GALMOSS 使用 PyTorch 框架支持的自动微分与优化器对模型参数进行梯度下降优化。由于 PyTorch 集成的自动微分在不需要进行显示求导的同时具有计算代价小、精度高、使用方便的特点，降低了用户修改或自定义 GALMOSS 程序的难度。GALMOSS 还支持参数拟合的范围限制。针对梯度下降无法量化参数拟合不确定性的问题，GALMOSS 提供了协方差矩阵和自举重采样两种方法，他们在速度与稳健性上各有优劣。

在 GALMOSS 的面亮度轮廓重拟合实验中，我们对其性能进行了评估。我们在 ~ 8,000 张 Sloan 数字巡天的星系图像上进行了单 Sersic 的拟合，并与 GALFIT 的拟合结果进行准确性的评估。评估结果发现，GALMOSS 与 GALFIT 的结果，除了 Sersic 指数因子 n 在 GALFIT 结果 > 6 的区域不太一致外，其他结果的一致性非常高。为了评估 GALMOSS 的速度，我们利用 Python 脚本模拟了 GALFIT 拟合相同数量的星系所需要的时间。实验结果发现，与 GALFIT 相比，GALMOSS 最高可以实现 6 倍的速度增幅。在具体的应用中，增幅的大小取决于样本大小与硬件（如 GPU、CPU）的性能。

目前，galmoss 已开源上传至 GitHub¹，Zenodo² 与 Python Package Index³ 三个平台，可进行便捷下载。代码文档已上传至 The readthedocs⁴，其中包括代码的总体介绍与使用示例等。附录A也包含了部分常用的使用场景示例，如单轮廓、双轮廓拟合、多轮廓组合建模与用户自定义轮廓等。

¹<https://github.com/Chenmi0619/GALMoss/>

²<https://doi.org/10.5281/zenodo.10654784>

³<https://pypi.org/project/galmoss/>

⁴<https://galmoss.readthedocs.io/en/latest/>

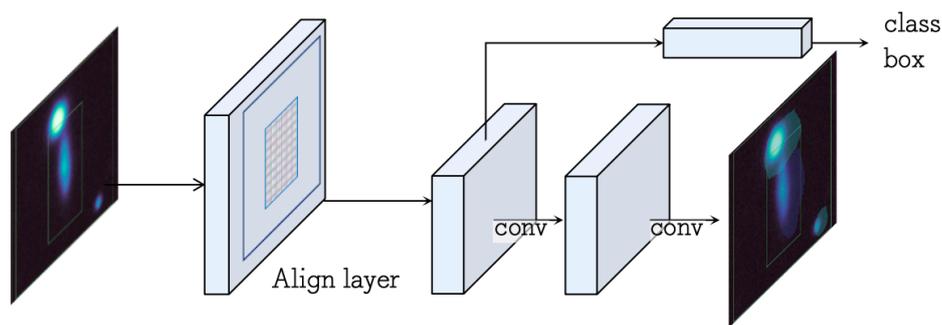


图 4.1 Mask-RCNN 的流程图。Mask-RCNN 在提取特征图之后，通过 Align layer 将特征图映射为相同大小。随后，该网络分为两个分支，第一个分支用于对此特征图进行分类，第二个分支用于产生掩膜图像。

4.2 讨论与未来展望

GALMOSS 的应用前景在于快速拟合大样本星系的面亮度轮廓，并根据所提取的星系特征展开后续的科学探究。比如，单成分的拟合能够很好地概括星系的总体特征，多成分的拟合可以对星系进行分解，并概括其各自子成分的特征。这些特征可用于探索星系形成与演化的路径，加深我们对宇宙演化的理解。

然而，面对这一需求，GALMOSS 仍然存在许多不足。在目前的大样本拟合工作中，用于拟合星系图像大小通常与星系的半光半径相关联。因而，大星系将拥有较大的星系图像大小，小星系则反之。但目前，GALMOSS 的并行化拟合流程要求星系图像裁剪的大小完全一致。在大星系的拟合中，这将导致图像裁剪的大小不足，并进而出现天空背景的错误估计；在小星系的拟合中，这将导致资源的浪费。在后续工作中，我们计划对星系裁剪的图像大小进行自适应，以实现更优的拟合。

其次，GALMOSS 的自动化程度依然有提升的空间。目前，除了基础的星系流量图像、点扩散函数图像与观测误差图像，GALMOSS 还要求用户提供拟合参数的初始值。在星系图像周围有强干扰源的情况下，GALMOSS 还需求准确的掩膜图像。后两者在望远镜二级数据发布时基本不会包括，仍需人工使用其他算法或软件进行介入。在后续工作中，我们计划利用 Mask-RCNN 与 ResNet50 实现自动的生成掩膜图像与生成初始值。相对准确的初始值能够定位到正确的局域坑周围，减少囿于错误局域最小值的概率，实现较为稳定的收敛。

如图4.1的流程图所示，Mask-RCNN 首先在星系流量图像上提取对应的特征图，随后通过 Align layer 将所提取的特征图映射为相同大小。在获取相同尺寸的特征图后，该网络分为两个分支，第一个分支用于对此特征图进行分类，第二个分支用于在特征图中产生更加细节的掩膜图像。在制作数据集的过程中，我们使用 128×128 的 DESI 巡天星系图像数据与其对应的 SExtractor 所产生的掩膜图像作为数据与标签，对 Mask-RCNN 网络进行训练。训练过程中，我们将背景作为第一类，中央星系作为第二类，干扰天体作为第三类。在使用训练集完成训练

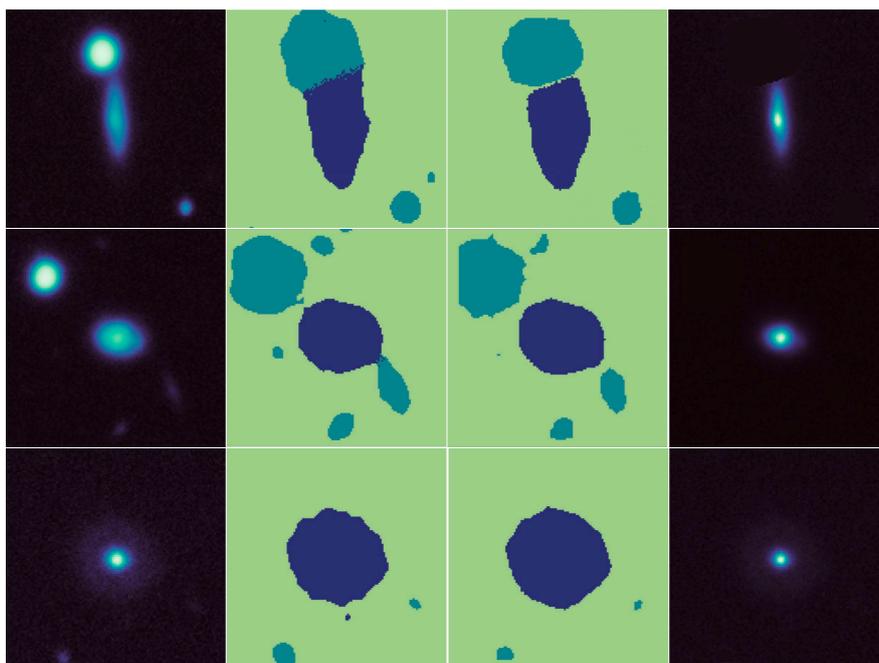


图 4.2 Mask-RCNN 的训练对比图。第一列为原图（为了凸显出中央星系所在位置，图像经过拉伸）；第二列为 SExtractor 所产生的掩膜图像；第三列为 Mask-RCNN 产生的掩膜图像，第四列为利用 Mask-RCNN 的掩膜图像去除干扰天体后的星系流量图像。

后，我们将模型应用于未参与训练过程的测试集图像中。如图4.2所示，我们的模型可以很好地将中央星系与干扰天体区分开，成功生成掩膜图像。

在利用 ResNet50 作为经典的卷积神经网络，常用于回归任务之中。单轮廓的参数回归由于参数量较小，ResNet50 可以在完成映射的基础上，减少过拟合的可能性。在制作数据集的过程中，我们使用经掩膜处理的 DESI 巡天星系图像数据与 galmass 所拟合的参数结果作为数据集，其图像大小同样为 128×128 。与 Mask-RCNN 的过程相类似，在使用训练集完成训练后，我们将模型应用于未参与训练过程的测试集图像中，并与 galmass 的结果相对比以测试其准确性。如图4.3所示，我们的模型所直接计算的结果与 galmass 所拟合的结果有较好地一致性。尽管部分参数的弥散相对较大，但神经网络的目的在于生成大致准确的初始参数，在输入 galmass 进行拟合之后，最终的结果最后依然可以收敛到最佳值。

神经网络最大的缺陷在于无法解决天文图像的域适应问题。不同的巡天虽然拍摄的天区相同，但由于观测深度、视宁度、分辨率的不同，神经网络的鲁棒性会大打折扣。因此，我们目前无法训练一个通用的网络，其有能力产生所有巡天与其对应波段的掩膜图像与初始参数。针对这一问题，我们计划对每一个巡天的对应波段进行针对性的训练与域迁移，并释放对应的网络权重。用户可以通过选择对应的网络权重，实现由星系图像数据集到 galmass 的全自动化拟合流程。将来我们将对这几个模块进行进一步的整合，进一步提升用户的使用感受。

然而，神经网络的权重依然依赖于相同的图像尺寸。在 128×128 大小的图

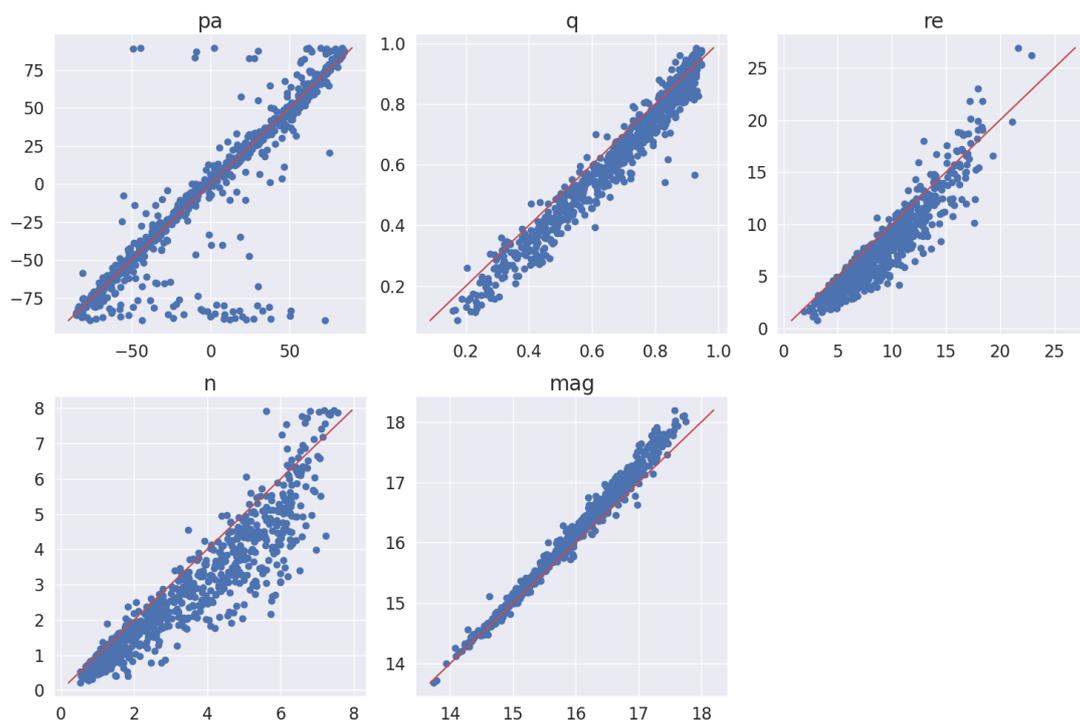


图 4.3

像上训练的网络权重，无法直接使用于其他大小的图像上。这与上文所提及的问题相类似，这些尺寸的适应性问题的都有待未来进行进一步的优化。

附录 A galmoss 使用示例

`galmoss` 灵活性强，支持多轮廓组合拟合以及用户自定义轮廓等多种用法。在本章节中，我们将就总体用法与细节设置做更加详尽的示例。在章节A.1中，我们将以 Sérsic 轮廓为例，介绍如何进行单轮廓拟合；在章节A.2中，我们将以核球、盘成分分解为例，介绍共享参数的双轮廓拟合；在章节A.3中，我们将介绍一个多轮廓建模的示例；在章节A.4中，我们将以更复杂的天空轮廓为例，介绍如何进行轮廓的用户自定义。本章节中所使用的图像数据可在 [GitHub](#)¹ 上进行下载。

A.1 单轮廓使用示例

在本节中，我们将介绍如何使用 `galmoss` 进行单轮廓拟合。在示例中，我们将使用 Sérsic 轮廓。

首先，我们需要加载所需要的 `galmoss` 包。

```
1 import Galmoss as gm
```

下一步，我们进行拟合参数与拟合轮廓的定义。由于此示例中并未包含多轮廓共用参数的要求，所以参数类的定义可以与轮廓类的定义（即参数类的传入）同时进行。在 `galmoss` 中，由于拟合参数均继承自共同的参数父类，因此在特殊参数（如方位角、半光半径、面亮度）的定义时，需要在数值传入之外进行额外的设置。在 `galmoss` 中，Sérsic 轮廓可选择是否对 `boxiness` 参数进行设置，在本示例中并未设置此参数。

```
1 # define parameter objects and profile
3 sersic = gm.lp.Sersic(
4     cen_x=gm.p(65.43),
5     cen_y=gm.p(64.95),
6     pa=gm.p(-81.06, angle=True),
7     axis_r=gm.p(0.64),
8     eff_r=gm.p(7.58, pix_scale=0.396),
9     ser_n=gm.p(1.53, log=True),
10    mag=gm.p(17.68, M0=22.5)
11 )
```

下一步，我们进行数据集的定义。数据集的定义包括星系的索引以及图像组（星系图像、误差图像、点扩散函数图像、以及掩膜图像）的路径。掩膜的索引

¹<https://github.com/Chenmi0619/GALMoss/>

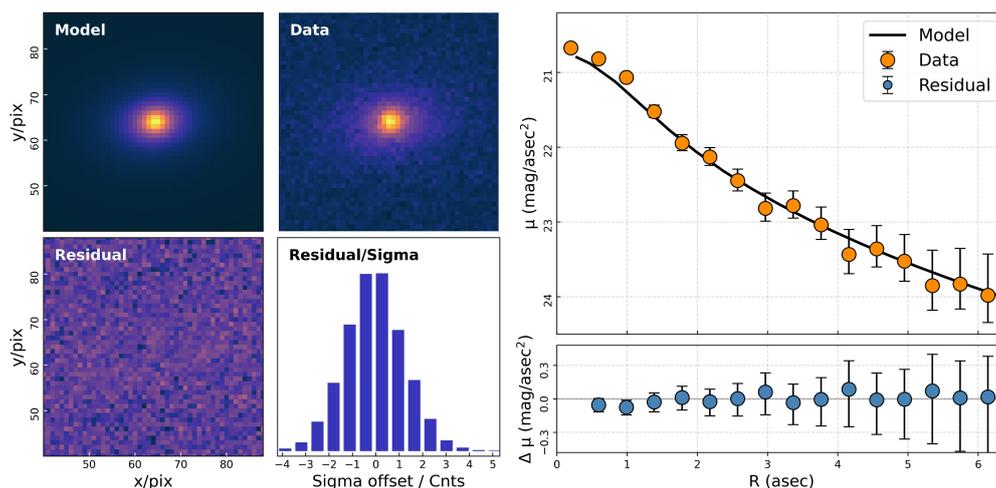


图 A.1 J162123.19+322056.4 星系 g 波段面亮度拟合的结果。左边为数据、模型与残差。右边为结果的一维投影。

默认为 1，如果含有非 1 的掩膜索引则需要手动传入。除此之外，数据集还需传入结果图像与拟合参数的保存路径。

```

1 dataset = gm.Dataset(
2     galaxy_index="J162123.19+322056.4",
3     image_path="./J162123.19+322056.4_image.fits",
4     sigma_path="./J162123.19+322056.4_sigma.fits",
5     psf_path="./J162123.19+322056.4_psf.fits",
6     mask_path="./J162123.19+322056.4_mask.fits"
7     mask_index=2,
8     img_block_path="./test_repo",
9     result_path="./test_repo"
10 )
11
dataset.define_profiles(sersic=sersic)

```

下一步，我们进行拟合参数的设置，如拟合批数量、最大迭代次数等。设置完成后即可进行拟合与拟合误差的计算。在此示例中，我们选取协方差矩阵的方法进行计算。

```

1 fitting = gm.Fitting(dataset=dataset,
2                     batch_size=1,
3                     iteration=1000)
4 fitting.fit()
5 fitting.uncertainty(method="covar_mat")

```

当拟合过程全部结束后，拟合参数与图像结果将保存至数据集定义时所传入的路径。图A.1展示了此示例的拟合结果。

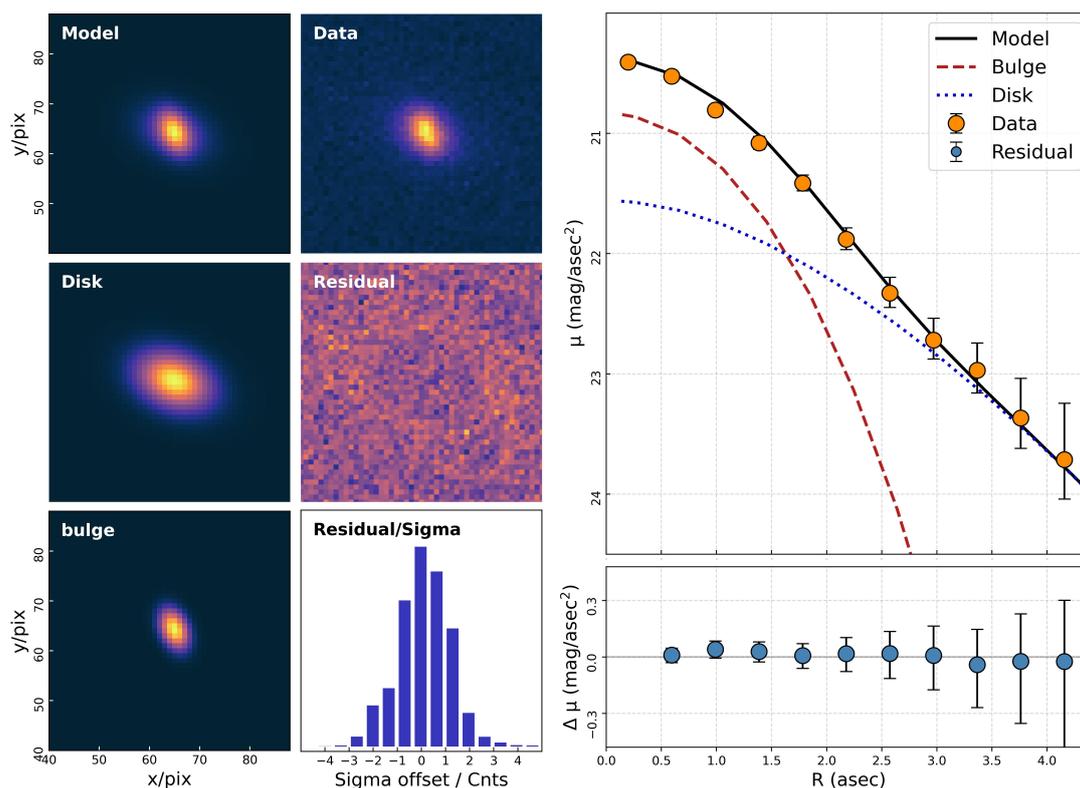


图 A.2 J100247.00+042559.8 星系 g 波段面亮度拟合的结果。左边为数据、模型与残差。右边为结果的一维投影。

A.2 双轮廓使用示例

在本节中，我们将介绍如何使用 galmass 进行双轮廓拟合。在示例中，我们将使用两个共享参数的 Sérsic 轮廓进行核球、盘成分分解。

```
import Galmass as gm
```

在导入 galmass 包后，我们进行拟合参数与拟合轮廓的定义。与单轮廓的拟合稍有不同的是，我们希望核球与盘成分应当具有相同的轮廓中心。因此，核球与盘两个轮廓需要导入同一个参数中心类。因此，参数中心类需要单独定义，无法与轮廓类同时定义。

```
1 xcen = gm.p(65.97)
  ycen = gm.p(65.30)
```

其他参数类则与单轮廓定义的步骤相同。

```
bulge = gm.lp.Sersic(
2     cen_x=xcen,
     cen_y=ycen,
4     pa=gm.p(58.70, angle=True),
     axis_r=gm.p(0.75),
```

```

6     eff_r=gm.p(4.09, pix_scale=0.396),
7     ser_n=gm.p(4),
8     mag=gm.p(17.97, M0=22.5)
9 )
10
11 disk = gm.lp.Sersic(
12     cen_x=xcen,
13     cen_y=ycen,
14     pa=gm.p(ini_value=58.70, angle=True),
15     axis_r=gm.p(0.75),
16     eff_r=gm.p(ini_value=4.09, pix_scale=0.396),
17     ser_n=gm.p(ini_value=1),
18     mag=gm.p(ini_value=17.97, M0=22.5)
19 )

```

数据集的定义、拟合参数的设置也与单轮廓拟合的步骤一致。在此示例中，我们使用自举重采样法进行拟合误差的计算。

```

dataset = gm.Data_Box(
2     galaxy_index="J100247.00+042559.8",
3     image_path="./J100247.00+042559.8_image.fits",
4     sigma_path="./J100247.00+042559.8_sigma.fits",
5     psf_path="./J100247.00+042559.8_psf.fits",
6     mask_path="./J100247.00+042559.8_mask.fits"
7     img_block_path="./test_repo",
8     result_path="./test_repo"
9 )
10
11 dataset.define_profiles(bulge=bulge, disk=disk)
12 fitting = gm.Fitting(dataset=dataset,
13                       batch_size=1,
14                       iteration=1000)
15 fitting.fit()
16 fitting.uncertainty(method="bstrap")

```

图A.2展示了此示例的拟合结果。

A.3 多轮廓建模

除了拟合需求之外，`galmass` 也可以根据所输入的参数数值进行模型图像的输出。在此示例中，我们将进行大量轮廓模型的组合，并对其模型图像的计算与输出。如图A.3与图A.4所示，我们在此组合中包含了一个拥有盘、核球、棒成分的中央星系，其由两个 Sérsic 轮廓与一个 Ferrer 轮廓组合而成。除此之外，组合中还包含了一个由单 Sérsic 轮廓定义的背景星系和两个由 Moffat 轮廓定义的点源。所有的面源轮廓都卷积了一个由高斯轮廓定义的点扩散函数图像。

```

import numpy as np
2 import Galmass as gm
import torch

```

首先，我们定义一个包含核球、盘与棒成分的中央星系。三个子成分共享轮廓中心。

```

1 # The galaxy center
xcen = gm.p(65.)
3 ycen = gm.p(64.)

5 # Define parameters and profiles
Bisk = gm.lp.Sersic(
7     cen_x=xcen,
     cen_y=ycen,
9     pa=gm.p(55., angle=True),
     axis_r=gm.p(0.44),
11    eff_r=gm.p(13, pix_scale=0.396),
     ser_n=gm.p(1),
13    mag=gm.p(17, M0=22.5),
     box=gm.p(0.1)
15 )

17 Bulge = gm.lp.Sersic(
     cen_x=xcen,
19    cen_y=ycen,
     pa=gm.p(0., angle=True),
21    axis_r=gm.p(1),
     eff_r=gm.p(6, pix_scale=0.396),
23    ser_n=gm.p(6),
     mag=gm.p(21.5, M0=22.5)
25 )

27 Bar = gm.lp.King(
     cen_x=xcen,
29    cen_y=ycen,
     pa=gm.p(-40., angle=True),
31    axis_r=gm.p(0.2),
     mag=gm.p(21.5, M0=22.5),
33    trunc_r=gm.p(5, pixScale=0.396),
     trunc_a=gm.p(0.5),
35    trunc_b=gm.p(1.9),
     box=gm.p(0.1)
37 )

```

其次，我们定义一个单轮廓的背景星系。

```
Sersic_s = gm.lp.Sersic(
```

```

2     cen_x=gm.p(45.43),
3     cen_y=gm.p(100.95),
4     pa=gm.p(-60., angle=True),
5     axis_r=gm.p(0.7),
6     eff_r=gm.p(13, pix_scale=0.396),
7     ser_n=gm.p(1.6),
8     mag=gm.p(17.5, M0=22.5),
9     box=gm.p(-0.6)
10 )

```

除了星系之外，我们使用高斯轮廓定义两个点源和点扩散函数图像。它们共享同一个 σ 数值。

```

# Two point - source
2 P1 = gm.lp.Gaussian(
3     cen_x= gm.p(75),
4     cen_y= gm.p(80),
5     pa=gm.p(0., angle=True),
6     axis_r=gm.p(1),
7     inten=gm.p(0.1),
8     fwhm=gm.p(1.2, pix_scale=0.396),
9 )
10
11 P2 = gm.lp.Gaussian(
12     cen_x=gm.p(88.),
13     cen_y=gm.p(40.),
14     pa=gm.p(0., angle=True),
15     axis_r=gm.p(1),
16     inten=gm.p(0.1),
17     fwhm=gm.p(1.2, pix_scale=0.396),
18 )
19
20 # The idealized PSF
21 PSF = gm.lp.Gaussian(
22     cen_x=gm.p(20),
23     cen_y=gm.p(20),
24     pa=gm.p(0., angle=True),
25     axis_r=gm.p(1),
26     inten=gm.p(1/(2*np.pi*((0.6/0.396)**2))),
27     fwhm=gm.p(1.2, pix_scale=0.396),
28 )

```

为了计算模型图像，我们需要手动定义一个网格，并调用轮廓类中的图像函数进行建模。图像函数中所使用的参数值将根据 `mode` 的模式来进行调用。在模型图像的计算中，我们使用 `initial_model`，即调用轮廓定义时所输入的参数值。总的模型图像（未经过卷积）将由这些子轮廓的图像叠加而成，其效果如图A.3所示。

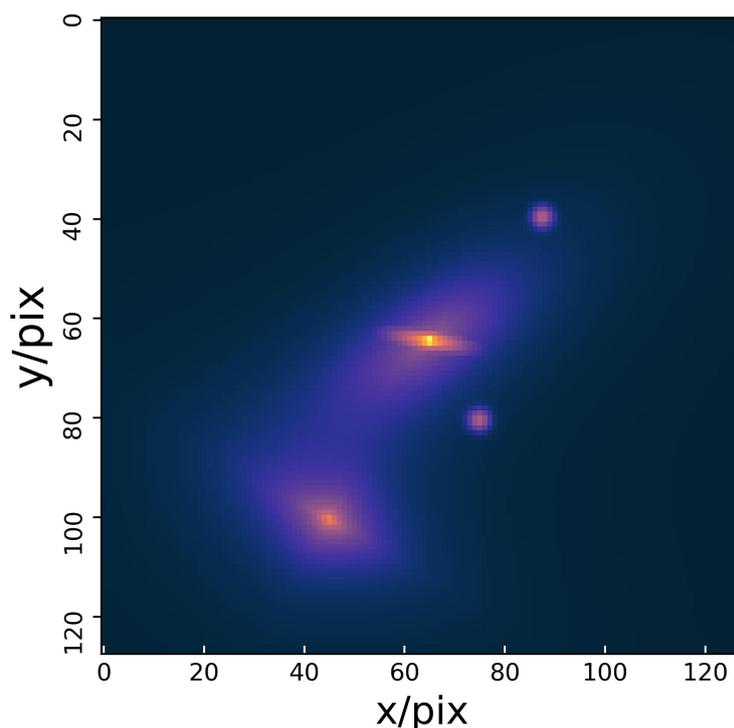


图 A.3 未经过 PSF 卷积的模型图像，包含中央星系、背景星系与两个点源。

```

1 x = torch.linspace(0.5, 128-0.5, 128)
2 y = torch.linspace(128-0.5, 0.5, 128)
3 grid = torch.meshgrid(y, x, indexing='ij')
4
5 bulge = Bulge.image_via_grid_from(grid, mode="initial_model")
6 disk = Disk.image_via_grid_from(grid, mode="initial_model")
7 bar = Bar.image_via_grid_from(grid, mode="initial_model")
8 sersic_s = Sersic_s.image_via_grid_from(grid, mode="initial_model")
9 p1 = P1.image_via_grid_from(grid, mode="initial_model")
10 p2 = P2.image_via_grid_from(grid, mode="initial_model")
11 result = bulge + disk + bar + sersic_s + p1 + p2

```

为了模拟视宁度的效应，我们使用一个较小的网格来模拟理想的点扩散函数图像，并卷积除了点源之外的其他模型图像。

```

1 #Grid generation
2 x_psf = torch.linspace(0.5, 40-0.5, 40)
3 y_psf = torch.linspace(40-0.5, 0.5, 40)
4 grid_psf = torch.meshgrid(x_psf, y_psf, indexing='ij')
5
6 # PSF image
7 psf = PSF.image_via_grid_from(grid_psf, mode="initial_model")
8
9 #Convoluting
10 result_conv = torch.squeeze(torch.nn.functional.conv2d((bulge+disk+bar+
11 sersic_s).reshape(1,1, 128,128), psf.reshape(1, 1, 40, 40), padding="same

```

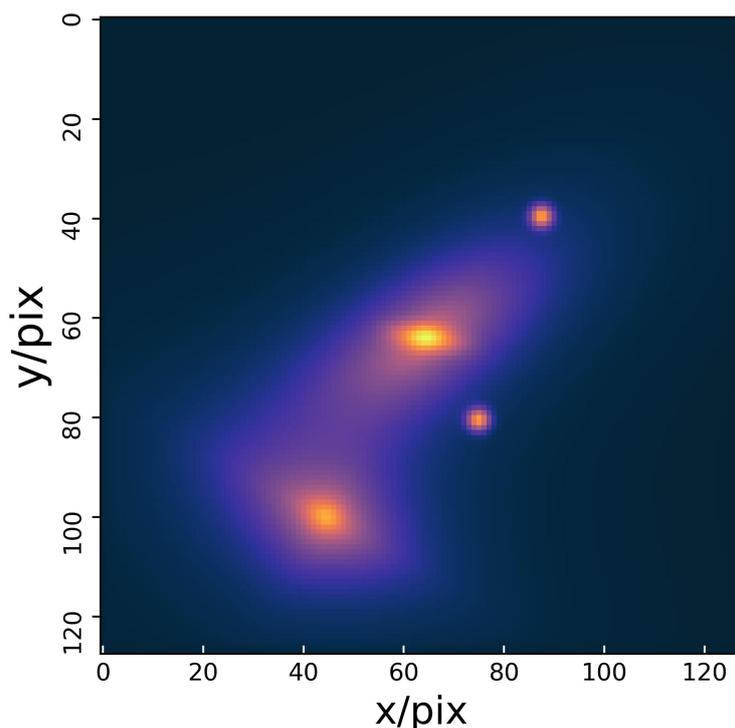


图 A.4 经过 PSF 卷积的模型图像，包含中央星系、背景星系与两个点源。

```
”, groups=1)) + p1 + p2
```

经卷积的模型图像如图A.4所示。

A.4 自定义轮廓示例

在本节中，我们将介绍如何使用 `galboss` 进行自定义新轮廓。在使用内置轮廓时，用户只需调用 `galboss` 中的对应类即可。但在使用自定义轮廓时，由于 `galboss` 中没有内置对应的类，用户需要进行新轮廓类的完整定义。

假设我们需要定义如下所示的背景轮廓，

$$I = I_0 + k_x(x - x_0) + k_y(y - y_0), \quad (\text{A.1})$$

其中 (x_0, y_0) 为图像的几何中心， I_0 为位于几何中心处的天空流量值，而 (k_x, k_y) 则描述了天空流量值如何随 x 轴与 y 轴变化。

我们将这个新轮廓类命名为 `NewSky`。这个类需要包括所需参数类的导入以及指导模型图像生成的函数。以下为几个注意事项：首先，所有的轮廓类需要继承 `LightProfile` 的基类，`galboss` 才能从而识别此为面亮度轮廓的类。其次，`__init__` 函数中需要加载所有所需的参数类，并设置为属性（如 `self.sky_0 = sky_0`）。以下为 `NewSky` 的完整定义。

```
import galmoss as gm
2
class NewSky(gm.LightProfile):
4     def __init__(self, sky_0, grad_x, grad_y):
6         super().__init__()
8         self.psf = False
10        self.sky_0 = sky_0
12        self.grad_x = grad_x
14        self.grad_y = grad_y
16
18        def image_via_grid_from(self,
                                grid,
                                mode="updating_model"):
        return (self.sky_0.value(mode)
                + self.grad_x.value(mode)
                * (grid[1] - (grid[0].shape[1] + 1)/2)
                + self.grad_y.value(mode)
                * (grid[0] - (grid[0].shape[0] + 1)/2))
```

轮廓的公式被包含在函数`image_via_grid_from`中。函数中所使用的参数值将根据 `mode` 的模式来进行调用，其默认模式为 `updating_model`，即每次优化迭代后，从优化器中取出的并调整相应维度的最新数值矩阵。

在新类的完整定义结束后，我们可以正常进入定义拟合参数与拟合轮廓的阶段。

```
1 sky = NewSky(sky_0=gm.p(0.3),
3             grad_x=gm.p(2),
4             grad_y=gm.p(3))
```

接下来的步骤则与前文完全一致。

参考文献

- Abraham R G, van den Bergh S. The morphological evolution of galaxies [J]. *Science*, 2001, 293 (5533): 1273-1278.
- Ahn C P, Alexandroff R, Prieto C A, et al. The ninth data release of the sloan digital sky survey: first spectroscopic data from the sdss-iii baryon oscillation spectroscopic survey [J]. *The Astrophysical Journal Supplement Series*, 2012, 203(2): 21.
- Allen P D, Driver S P, Graham A W, et al. The millennium galaxy catalogue: bulge–disc decomposition of 10 095 nearby galaxies [J]. *Monthly Notices of the Royal Astronomical Society*, 2006, 371(1): 2-18.
- Aragon-Calvo M A, Carvajal J. Self-supervised learning with physics-aware neural networks–i. galaxy model fitting [J]. *Monthly Notices of the Royal Astronomical Society*, 2020, 498(3): 3713-3719.
- Athanassoula E, Morin S, Wozniak H, et al. The shape of bars in early-type barred galaxies [J]. *Monthly Notices of the Royal Astronomical Society (ISSN 0035-8711)*, vol. 245, July 1, 1990, p. 130-139., 1990, 245: 130-139.
- Banerji M, Lahav O, Lintott C J, et al. Galaxy zoo: reproducing galaxy morphologies via machine learning [J]. *Monthly Notices of the Royal Astronomical Society*, 2010, 406(1): 342-353.
- Baydin A G, Pearlmutter B A, Radul A A, et al. Automatic differentiation in machine learning: a survey [J]. *Journal of machine learning research*, 2018, 18(153): 1-43.
- Bernardi M, Meert A, Vikram V, et al. Systematic effects on the size–luminosity relations of early- and late-type galaxies: dependence on model fitting and morphology [J]. *Monthly Notices of the Royal Astronomical Society*, 2014, 443(1): 874-897.
- Bershady M A, Jangren A, Conselice C J. Structural and photometric classification of galaxies. i. calibration based on a nearby galaxy sample [J]. *The Astronomical Journal*, 2000, 119(6): 2645.
- Binggeli B, Cameron L. Dwarf galaxies in the virgo cluster. i-the systematic photometric properties of early-type dwarfs [J]. *Astronomy and Astrophysics (ISSN 0004-6361)*, vol. 252, no. 1, Dec. 1991, p. 27-52., 1991, 252: 27-52.
- Binggeli B, Sandage A, Tarenghi M. Studies of the virgo cluster. i-photometry of 109 galaxies near the cluster center to serve as standards [J]. *Astronomical Journal (ISSN 0004-6256)*, vol. 89, Jan. 1984, p. 64-82., 1984, 89: 64-82.
- Bothun G D, Mould J R, Caldwell N, et al. Comparative photometric parameters of dwarf irregular and elliptical galaxies in the virgo cluster-two different classes of dwarf galaxies? [J]. *Astronomical Journal (ISSN 0004-6256)*, vol. 92, Nov. 1986, p. 1007-1019., 1986, 92: 1007-1019.
- Bretonnière H, Boucaud A, Lanusse F, et al. Euclid preparation-xiii. forecasts for galaxy morphology with the euclid survey using deep generative models [J]. *Astronomy & Astrophysics*, 2022, 657: A90.
- Burkert A. The structure and dark halo core properties of dwarf spheroidal galaxies [J]. *The Astrophysical Journal*, 2015, 808(2): 158.
- Burstein D. Structure and origin of s0 galaxies. ii-disk-to-bulge ratios [J]. *Astrophysical Journal*, Part 1, vol. 234, Dec. 1, 1979, p. 435-447., 1979, 234: 435-447.
- Byun Y, Freeman K. Two-dimensional decomposition of bulge and disk [J]. *Astrophysical Journal* v. 448, p. 563, 1995, 448: 563.

- Caldwell N. Structure and stellar content of dwarf elliptical galaxies [J]. *Astronomical Journal* (ISSN 0004-6256), vol. 88, June 1983, p. 804-812., 1983, 88: 804-812.
- Caon N, Capaccioli M, D'Onofrio M. On the shape of the light profiles of early-type galaxies [J]. *Monthly Notices of the Royal Astronomical Society*, 1993, 265(4): 1013-1021.
- Capaccioli M. Structure and dynamics of elliptical galaxies [C]//IAU Symp 127: volume 47. D. Reidel Publ. Co., 1987.
- Capaccioli M. The world of galaxies, ed. hg corwin, l. bottinelli [M]. Berlin: Springer-Verlag, 1989.
- Capaccioli M, Caon N. On the lack of a simple relation between r_e and μ_e for early-type galaxies [J]. *Monthly Notices of the Royal Astronomical Society*, 1991, 248(3): 523-527.
- Capaccioli M, Held E V, Nieto J L. The lenticular ngc 3115: A standard for galaxy photometry [C]//New Aspects of Galaxy Photometry: Proceedings of the Specialized Meeting of the Eight IAU European Regional Astronomy Meeting Toulouse, September 17–21, 1984. Springer, 1985: 265-268.
- Cauchy A, et al. Méthode générale pour la résolution des systemes d' équations simultanées [J]. *Comp. Rend. Sci. Paris*, 1847, 25(1847): 536-538.
- Cava A, Schaerer D, Richard J, et al. The nature of giant clumps in distant galaxies probed by the anatomy of the cosmic snake [J]. *Nature Astronomy*, 2018, 2(1): 76-82.
- Cibinel A, Le Floc' h E, Perret V, et al. A physical approach to the identification of high-z mergers: morphological classification in the stellar mass domain [J]. *The Astrophysical Journal*, 2015, 805 (2): 181.
- Collaboration S. The seventeenth data release of the sloan digital sky surveys [J]. *The Astrophysical Journal Supplement Series*, 2022, 259(35): 39pp.
- Conroy C. Modeling the panchromatic spectral energy distributions of galaxies [J]. *Annual Review of Astronomy and Astrophysics*, 2013, 51: 393-455.
- Conselice C J. The evolution of starburst galaxies [J]. arXiv preprint astro-ph/0012454, 2000.
- Cowles M K, Carlin B P. Markov chain monte carlo convergence diagnostics: a comparative review [J]. *Journal of the American statistical Association*, 1996, 91(434): 883-904.
- Dayal P, Ferrara A. Early galaxy formation and its large-scale effects [J]. *Physics Reports*, 2018, 780: 1-64.
- de Jong R S. Near-infrared and optical broadband surface photometry of 86 face-on disk dominated galaxies. iii. the statistics of disk and bulge parameters [J]. arXiv preprint astro-ph/9601005, 1996.
- De Souza R, Gadotti D A, Dos Anjos S. Budda: a new two-dimensional bulge/disk decomposition code for detailed structural analysis of galaxies [J]. *The Astrophysical Journal Supplement Series*, 2004, 153(2): 411.
- de Vaucouleurs G. Recherches sur les nebuleuses extragalactiques [J]. *Annales d'Astrophysique*, Vol. 11, p. 247, 1948, 11: 247.
- de Vaucouleurs G. Photoelectric photometry of the andromeda nebula in the ubv system [J]. *Astrophysical Journal*, Vol. 128, p. 465, 1958, 128: 465.
- De Vaucouleurs G. Classification and morphology of external galaxies [M]//Astrophysik iv: Sternsysteme/astrophysics iv: Stellar systems. Springer, 1959: 275-310.
- de Zeeuw T, Franx M. Structure and dynamics of elliptical galaxies [J]. *Annual review of Astronomy and Astrophysics*, 1991, 29(1): 239-274.

- Domínguez Sánchez H, Huertas-Company M, Bernardi M, et al. Improving galaxy morphologies for sdss with deep learning [J]. *Monthly Notices of the Royal Astronomical Society*, 2018, 476(3): 3661-3676.
- D’Onofrio M, Capaccioli M, Caon N. On the shape of the light profiles of early type galaxies—ii. the diagram [J]. *Monthly Notices of the Royal Astronomical Society*, 1994, 271(3): 523-529.
- Drew P M, Casey C M, Cooray A, et al. Three dusty star-forming galaxies at $z \sim 1.5$: Mergers and disks on the main sequence [J]. *The Astrophysical Journal*, 2020, 892(2): 104.
- Dullo B T, Graham A W. Sizing up partially depleted galaxy cores [J]. *The Astrophysical Journal*, 2012, 755(2): 163.
- Efron B. Bootstrap methods: another look at the jackknife [M]//Breakthroughs in statistics: Methodology and distribution. Springer, 1992: 569-593.
- Erwin P. Imfit: a fast, flexible new program for astronomical image fitting [J]. *The Astrophysical Journal*, 2015, 799(2): 226.
- Erwin P, Pohlen M, Beckman J E. The outer disks of early-type galaxies. i. surface-brightness profiles of barred galaxies [J]. *The Astronomical Journal*, 2007, 135(1): 20.
- Faber S, Lin D. Is there nonluminous matter in dwarf spheroidal galaxies [J]. *Astrophysical Journal*, Part 2-Letters to the Editor, vol. 266, Mar. 1, 1983, p. L17-L20., 1983, 266: L17-L20.
- Ferguson H C, Dickinson M, Giavalisco M, et al. The size evolution of high-redshift galaxies [J]. *The Astrophysical Journal*, 2004, 600(2): L107.
- Ferrers N M. An elementary treatise on spherical harmonics and subjects connected with them [M]. Macmillan and Company, 1877.
- Fischer J, Domínguez Sánchez H, Bernardi M. Sdss-iv manga pymorph photometric and deep learning morphological catalogues and implications for bulge properties and stellar angular momentum [J]. *Monthly Notices of the Royal Astronomical Society*, 2019, 483(2): 2057-2077.
- Foex G, Soucail G, Pelló R, et al. The role of environment in the morphological transformation of galaxies in 9 rich intermediate redshift clusters [J]. *Astronomy & Astrophysics*, 2009, 505(1): 83-96.
- Foreman-Mackey D, Hogg D W, Lang D, et al. emcee: the mcmc hammer [J]. *Publications of the Astronomical Society of the Pacific*, 2013, 125(925): 306.
- Freeman K C. On the disks of spiral and s0 galaxies [J]. *Astrophysical Journal*, vol. 160, p. 811, 1970, 160: 811.
- Gao H, Ho L C. An optimal strategy for accurate bulge-to-disk decomposition of disk galaxies [J]. *The Astrophysical Journal*, 2017, 845(2): 114.
- Gauss C F. Bestimmung der genauigkeit der beobachtungen [J]. *Ibidem*, 1816: 129-138.
- Gauss C F. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*: volume 7 [M]. FA Perthes, 1877.
- Gavin H P. The levenberg-marquardt algorithm for nonlinear least squares curve-fitting problems [J]. Department of Civil and Environmental Engineering, Duke University Durham, NC USA, 2022.
- George A, Damjanov I, Sawicki M, et al. Two rest-frame wavelength measurements of galaxy sizes at $z < 1$: the evolutionary effects of emerging bulges and quenched newcomers [J]. *Monthly Notices of the Royal Astronomical Society*, 2024, 528(3): 4797-4828.
- George K. Structural analysis of star-forming blue early-type galaxies-merger-driven star formation in elliptical galaxies [J]. *Astronomy & Astrophysics*, 2017, 598: A45.

- Ghosh A, Urry C M, Rau A, et al. Gampen: A machine-learning framework for estimating bayesian posteriors of galaxy morphological parameters [J]. *The Astrophysical Journal*, 2022, 935(2): 138.
- Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [J]. *Advances in neural information processing systems*, 2014, 27.
- Governato F, Baugh C, Frenk C, et al. The seeds of rich galaxy clusters in the universe [J]. *Nature*, 1998, 392(6674): 359-361.
- Graham A W, Worley C C. Inclination-and dust-corrected galaxy parameters: bulge-to-disc ratios and size–luminosity relations [J]. *Monthly Notices of the Royal Astronomical Society*, 2008, 388 (4): 1708-1728.
- Graham A W, Erwin P, Trujillo I, et al. A new empirical model for the structural analysis of early-type galaxies, and a critical review of the nuker model [J]. *The Astronomical Journal*, 2003, 125 (6): 2951.
- Graham A W, Merritt D, Moore B, et al. Empirical models for dark matter halos. ii. inner profile slopes, dynamical profiles, and ρ/σ^3 [J]. *The Astronomical Journal*, 2006, 132(6): 2701.
- Grillmair C J, Faber S, Lauer T R, et al. The nuclear regions of ngc 3311 and ngc 7768 imaged with the hubble space telescope planetary camera [J]. *The Astronomical Journal*, vol. 108, no. 1, p. 102-110, 1994, 108: 102-110.
- Grover H, Bait O, Wadadekar Y, et al. Predicting bulge to total luminosity ratio of galaxies using deep learning [J]. *Monthly Notices of the Royal Astronomical Society*, 2021, 506(3): 3313-3322.
- Hammersley J. Monte carlo methods [M]. Springer Science & Business Media, 2013.
- Hart R E, Bamford S P, Casteels K R, et al. Galaxy zoo: star formation versus spiral arm number [J]. *Monthly Notices of the Royal Astronomical Society*, 2017, 468(2): 1850-1863.
- Hastings W K. Monte carlo sampling methods using markov chains and their applications [J]. 1970.
- Häußler B, Bamford S P, Vika M, et al. Megamorph–multiwavelength measurement of galaxy structure: complete sérsic profile information from modern surveys [J]. *Monthly Notices of the Royal Astronomical Society*, 2013, 430(1): 330-369.
- He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- Hochreiter S, Schmidhuber J. Long short-term memory [J]. *Neural computation*, 1997, 9(8): 1735-1780.
- Hodge P W. The fornax dwarf galaxy. i. the globulars clusters. [J]. *Astronomical Journal*, Vol. 66, p. 83-84 (1961), 1961, 66: 83-84.
- Hodge P W. The fornax dwarf galaxy. ii. the distribution of stars [J]. *Astronomical Journal*, Vol. 66, p. 249 (1961), 1961, 66: 249.
- Hodge P W. Dwarf galaxies [J]. *Annual Review of Astronomy and Astrophysics*, 1971, 9(1): 35-66.
- Hogg D W, Bovy J, Lang D. Data analysis recipes: Fitting a model to data [J]. arXiv preprint arXiv:1008.4686, 2010.
- Hubble E. A relation between distance and radial velocity among extra-galactic nebulae [J]. *Proceedings of the national academy of sciences*, 1929, 15(3): 168-173.
- Hubble E P. Extragalactic nebulae. [J]. *Astrophysical Journal*, 64, 321-369 (1926), 1926, 64.
- Ichikawa S I, Wakamatsu K I, Okamura S. Surface photometry of dwarf elliptical galaxies in the virgo cluster [J]. *Astrophysical Journal Supplement Series (ISSN 0067-0049)*, vol. 60, Feb. 1986, p. 475-506. Sponsorship: Ministry of Education, Science, and Culture of Japan., 1986, 60: 475-506.

- Jeans J H. The stability of a spherical nebula [J]. Proceedings of the Royal Society of London, 1901, 68(442-450): 454-455.
- Jerjen H, Freeman K C, Binggeli B. Surface brightness fluctuation distances to dwarf elliptical galaxies in the sculptor group [J]. The Astronomical Journal, 1998, 116(6): 2873.
- Kennedy R, Bamford S P, Häußler B, et al. Galaxy and mass assembly (gama): understanding the wavelength dependence of galaxy structure with bulge-disc decompositions [J]. Monthly Notices of the Royal Astronomical Society, 2016, 460(4): 3458-3471.
- Kent S M. Ccd surface photometry of field galaxies. ii-bulge/disk decompositions [J]. Astrophysical Journal Supplement Series (ISSN 0067-0049), vol. 59, Oct. 1985, p. 115-159., 1985, 59: 115-159.
- Kim J h, Wise J H, Alvarez M A, et al. Galaxy formation with self-consistently modeled stars and massive black holes. i. feedback-regulated star formation and black hole growth [J]. The Astrophysical Journal, 2011, 738(1): 54.
- Kinakh V, Belousov Y, Quétant G, et al. Hubble meets webb: Image-to-image translation in astronomy [J]. Sensors, 2024, 24(4): 1151.
- King I. The structure of star clusters. i. an empirical density law [J]. Astronomical Journal, Vol. 67, p. 471 (1962), 1962, 67: 471.
- King I R. The structure of star clusters. iii. some simple dynamical models [J]. Astronomical Journal, Vol. 71, p. 64 (1966), 1966, 71: 64.
- King I R, Minkowski R. Mass-luminosity ratios and sizes of giant elliptical galaxies [C]//External Galaxies and Quasi-Stellar Objects: volume 44. 1972: 87.
- Kingma D P, Welling M. Auto-encoding variational bayes [J]. arXiv preprint arXiv:1312.6114, 2013.
- Kingma D P, Ba J A, Adam J. A method for stochastic optimization. arxiv 2014 [J]. arXiv preprint arXiv:1412.6980, 2020, 106: 1.
- Kormendy J, Dressler A, Byun Y, et al. Dwarf galaxies ed g [J]. Meylan and P. Prugniel (Garching: ESO), 1994, 147.
- Kormendy J. Brightness distributions in compact and normal galaxies. ii-structure parameters of the spheroidal component [J]. Astrophysical Journal, Part 1, vol. 218, Dec. 1, 1977, p. 333-346. Research supported by the California Institute of Technology and National Research Council of Canada., 1977, 218: 333-346.
- Kormendy J, Kennicutt Jr R C. Secular evolution and the formation of pseudobulges in disk galaxies [J]. Annu. Rev. Astron. Astrophys., 2004, 42: 603-683.
- Kremer J, Stensbo-Smidt K, Gieseke F, et al. Big universe, big data: machine learning and image analysis for astronomy [J]. IEEE Intelligent Systems, 2017, 32(2): 16-22.
- Laine J, Laurikainen E, Salo H, et al. Morphology and environment of galaxies with disc breaks in the s4g and nirs0s [J]. Monthly Notices of the Royal Astronomical Society, 2014, 441(3): 1992-2012.
- Lange R, Driver S P, Robotham A S, et al. Galaxy and mass assembly (gama): mass-size relations of $z < 0.1$ galaxies subdivided by sérsic index, colour and morphology [J]. Monthly Notices of the Royal Astronomical Society, 2015, 447(3): 2603-2630.
- Larson R B. Numerical calculations of the dynamics of a collapsing proto-star [J]. Monthly Notices of the Royal Astronomical Society, 1969, 145(3): 271-295.
- Lauer T R. The cores of elliptical galaxies [J]. Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 292, May 1, 1985, p. 104-121., 1985, 292: 104-121.

- Lauer T R, Ajhar E A, Byun Y I, et al. The centers of early-type galaxies with hst. i. an observational survey [J]. *Astronomical Journal* v. 110, p. 2622, 1995, 110: 2622.
- Levenberg K. A method for the solution of certain non-linear problems in least squares [J]. *Quarterly of applied mathematics*, 1944, 2(2): 164-168.
- Li R, Napolitano N, Roy N, et al. Galaxy light profile convolutional neural networks (galnets). i. fast and accurate structural parameters for billion-galaxy samples [J]. *The Astrophysical Journal*, 2022, 929(2): 152.
- Liu C, Hao L, Wang H, et al. The morphological transformation and the quenching of galaxies [J]. *The Astrophysical Journal*, 2019, 878(1): 69.
- Lotz J M, Primack J, Madau P. A new nonparametric approach to galaxy morphological classification [J]. *The Astronomical Journal*, 2004, 128(1): 163.
- Lotz J M, Davis M, Faber S, et al. The evolution of galaxy mergers and morphology at $z < 1.2$ in the extended groth strip [J]. *The Astrophysical Journal*, 2008, 672(1): 177.
- Majewski S, Forestell A, Ostheimer J, et al. The extended shapes of galactic satellites [M]//*The Shapes of Galaxies and their Dark Halos*. World Scientific, 2002: 214-221.
- Marquardt D W. An algorithm for least-squares estimation of nonlinear parameters [J]. *Journal of the society for Industrial and Applied Mathematics*, 1963, 11(2): 431-441.
- Martín-Navarro I, Brodie J P, Romanowsky A J, et al. Black-hole-regulated star formation in massive galaxies [J]. *Nature*, 2018, 553(7688): 307-309.
- Meert A, Vikram V, Bernardi M. A catalogue of 2d photometric decompositions in the sdss-dr7 spectroscopic main galaxy sample: preferred models and systematics [J]. *Monthly Notices of the Royal Astronomical Society*, 2015, 446(4): 3943-3974.
- Metropolis N, Rosenbluth A W, Rosenbluth M N, et al. Equation of state calculations by fast computing machines [J]. *The journal of chemical physics*, 1953, 21(6): 1087-1092.
- Meyn S P, Tweedie R L. Computable bounds for geometric convergence rates of markov chains [J]. *The Annals of Applied Probability*, 1994: 981-1011.
- Mineshige S, Nomura H, Hirose M, et al. Black hole disk accretion in supernovae [J]. *The Astrophysical Journal*, 1997, 489(1): 227.
- Miranda L J. Pyswarms: a research toolkit for particle swarm optimization in python [J]. *Journal of Open Source Software*, 2018, 3(21): 433.
- Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention [J]. *Advances in neural information processing systems*, 2014, 27.
- Modi C, Lanusse F, Seljak U. Flowpm: Distributed tensorflow implementation of the fastpm cosmological n-body solver [J]. *Astronomy and Computing*, 2021, 37: 100505.
- Moffat A. A theoretical investigation of focal stellar images in the photographic emulsion and application to photographic photometry [J]. *Astronomy and Astrophysics*, 1969, 3: 455.
- Mosleh M, Hosseinnejad S, Hosseini-ShahiSavandi S Z, et al. Galaxy sizes since $z = 2$ from the perspective of stellar mass distribution within galaxies [J]. *The Astrophysical Journal*, 2020, 905(2): 170.
- Nair P B, Abraham R G. A catalog of detailed visual morphological classifications for 14,034 galaxies in the sloan digital sky survey [J]. *The Astrophysical Journal Supplement Series*, 2010, 186(2): 427.
- Nelder J A, Mead R. A simplex method for function minimization [J]. *The computer journal*, 1965, 7(4): 308-313.

- Nightingale J, Amvrosiadis A, Hayes R G, et al. Pyautogalaxy: Open-source multiwavelength galaxy structure & morphology [J]. *Journal of Open Source Software*, 2023, 8(81).
- Oswalt T D, Keel W C. Planets, stars and stellar systems: Volume 6: Extragalactic astronomy and cosmology [M]. 2013.
- Pacifici C, Iyer K G, Mobasher B, et al. The art of measuring physical parameters in galaxies: a critical assessment of spectral energy distribution fitting techniques [J]. *The Astrophysical Journal*, 2023, 944(2): 141.
- Park C, Hwang H S. Interactions of galaxies in the galaxy cluster environment [J]. *The Astrophysical Journal*, 2009, 699(2): 1595.
- Pasha I. pysersic: A python package for determining galaxy structural properties via bayesian inference, accelerated with jax [J]. *Monthly Notices of the Royal Astronomical Society*, 2022.
- Paulino-Afonso A, Sobral D, Buitrago F, et al. The structural and size evolution of star-forming galaxies over the last 11 gyr [J]. *Monthly Notices of the Royal Astronomical Society*, 2016, 465(3): 2717-2733.
- Pawitan Y. In all likelihood: statistical modelling and inference using likelihood [M]. Oxford University Press, 2001.
- Peng C Y. Galfit user's manual [J]. Retrieved from, 2003.
- Peng C Y, Ho L C, Impey C D, et al. Detailed structural decomposition of galaxy images [J]. *The Astronomical Journal*, 2002, 124(1): 266.
- Peng C Y, Ho L C, Impey C D, et al. Detailed decomposition of galaxy images. ii. beyond axisymmetric models [J]. *The Astronomical Journal*, 2010, 139(6): 2097.
- Plummer H C. On the problem of distribution in globular star clusters [J]. *Monthly Notices of the Royal Astronomical Society*, Vol. 71, p. 460-470, 1911, 71: 460-470.
- Pohlen M, Trujillo I. The structure of galactic disks-studying late-type spiral galaxies using sdss [J]. *Astronomy & Astrophysics*, 2006, 454(3): 759-772.
- Qiu C, Napolitano N R, Li R, et al. Galaxy light profile neural networks (galnets). ii. bulge-disc decomposition in optical space-based observations [J]. arXiv preprint arXiv:2306.05909, 2023.
- Rest A, van den Bosch F C, Jaffe W, et al. Wfpc2 images of the central regions of early-type galaxies. i. the data [J]. *The Astronomical Journal*, 2001, 121(5): 2431.
- Reynolds J. The light curve of the andromeda nebula (ngc 224) [J]. *Monthly Notices of the Royal Astronomical Society*, Vol. 74, p. 132-136, 1913, 74: 132-136.
- Rigamonti F, Dotti M, Covino S, et al. Decomposing galaxies with bang: an automated morphokinematic decomposition of the sdss-dr17 manga survey [J]. *Monthly Notices of the Royal Astronomical Society*, 2023, 525(1): 1008-1022.
- Robotham A, Taranu D, Tobar R, et al. Profit: Bayesian profile fitting of galaxy images [J]. *Monthly Notices of the Royal Astronomical Society*, 2017, 466(2): 1513-1541.
- Robotham A, Davies L, Driver S, et al. Profound: source extraction and application to modern survey data [J]. *Monthly Notices of the Royal Astronomical Society*, 2018, 476(3): 3137-3159.
- Rubinstein R Y, Kroese D P. Simulation and the monte carlo method [M]. John Wiley & Sons, 2016.
- Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [J]. *nature*, 1986, 323(6088): 533-536.
- Savorgnan G A, Graham A W. Supermassive black holes and their host spheroids. i. disassembling galaxies [J]. *The Astrophysical Journal Supplement Series*, 2016, 222(1): 10.
- Schade D, Lilly S, Crampton D, et al. Canada-france redshift survey: Hubble space telescope imaging of high-redshift field galaxies [J]. *The Astrophysical Journal*, 1995, 451(1): L1.

- Schoenholz S, Cubuk E D. Jax md: a framework for differentiable physics [J]. *Advances in Neural Information Processing Systems*, 2020, 33: 11428-11441.
- Schwarzschild M. A numerical model for a triaxial stellar system in dynamical equilibrium [J]. *Astrophysical Journal, Part 1*, vol. 232, Aug. 15, 1979, p. 236-247., 1979, 232: 236-247.
- Send U. Surface photometry of the spiral galaxy ngc 4501 [J]. *Astronomy and Astrophysics*, vol. 112, no. 2, Aug. 1982, p. 235-240., 1982, 112: 235-240.
- Sérsic J. Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy [J]. *Boletín de la Asociacion Argentina de Astronomia La Plata Argentina*, 1963, 6: 41-43.
- Shaw M A, Gilmore G. The luminosity distributions of edge-on spiral galaxies–i. a two-dimensional model and its application to ngc 891 and 4565 [J]. *Monthly Notices of the Royal Astronomical Society*, 1989, 237(4): 903-927.
- Shen M t, Yin J, Hao L, et al. Comparison of photometric and morphological decompositions catalogues [J]. *Progress in Astronomy*, 2022, 40: 228-249.
- Shen S, Mo H, White S D, et al. The size distribution of galaxies in the sloan digital sky survey [J]. *Monthly Notices of the Royal Astronomical Society*, 2003, 343(3): 978-994.
- Simard L. Gim2d: an iraf package for the quantitative morphology analysis of distant galaxies [C]// *Astronomical Data Analysis Software and Systems VII: volume 145*. 1998: 108.
- Simard L, Willmer C N, Vogt N P, et al. The deep groth strip survey. ii. hubble space telescope structural parameters of galaxies in the groth strip [J]. *The Astrophysical Journal Supplement Series*, 2002, 142(1): 1.
- Simard L, Mendel J T, Patton D R, et al. A catalog of bulge+ disk decompositions and updated photometry for 1.12 million galaxies in the sloan digital sky survey [J]. *The Astrophysical Journal Supplement Series*, 2011, 196(1): 11.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv preprint arXiv:1409.1556*, 2014.
- Sohl-Dickstein J, Weiss E, Maheswaranathan N, et al. Deep unsupervised learning using nonequilibrium thermodynamics [C]// *International conference on machine learning*. PMLR, 2015: 2256-2265.
- Speagle J S. dynesty: a dynamic nested sampling package for estimating bayesian posteriors and evidences [J]. *Monthly Notices of the Royal Astronomical Society*, 2020, 493(3): 3132-3158.
- Springel V, Hernquist L. The history of star formation in a λ cold dark matter universe [J]. *Monthly Notices of the Royal Astronomical Society*, 2003, 339(2): 312-334.
- Stone C J, Courteau S, Cuillandre J C, et al. Astrophot: Fitting everything everywhere all at once in astronomical images [J]. *Monthly Notices of the Royal Astronomical Society*, 2023: stad2477.
- Storn R, Price K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces [J]. *Journal of global optimization*, 1997, 11: 341-359.
- Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks [C]// *International conference on machine learning*. PMLR, 2019: 6105-6114.
- Thompson R, Davé R, Huang S, et al. Identifying mergers using non-parametric morphological classification at high redshifts [J]. *arXiv preprint arXiv:1508.01851*, 2015.
- Tortorelli L, Mercurio A. Morphofit: An automated galaxy structural parameters fitting package [J]. *Frontiers in Astronomy and Space Sciences*, 2023, 10: 989443.

- Toyouchi D, Hosokawa T, Sugimura K, et al. Gaseous dynamical friction under radiative feedback: do intermediate-mass black holes speed up or down? [J]. *Monthly Notices of the Royal Astronomical Society*, 2020, 496(2): 1909-1921.
- Tripathi A, Panwar N, Sharma S, et al. Photometric and kinematic study of the open cluster ngc 1027 [J]. *arXiv preprint arXiv:2304.05762*, 2023.
- Trujillo I, Graham A W, Caon N. On the estimation of galaxy structural parameters: the sérsic model [J]. *Monthly Notices of the Royal Astronomical Society*, 2001, 326(3): 869-876.
- Trujillo I, Erwin P, Ramos A A, et al. Evidence for a new elliptical-galaxy paradigm: Sérsic and core galaxies [J]. *The Astronomical Journal*, 2004, 127(4): 1917.
- Tuccillo D, Huertas-Company M, Decencière E, et al. Deep learning for galaxy surface brightness profile fitting [J]. *Monthly Notices of the Royal Astronomical Society*, 2018, 475(1): 894-909.
- Van Dokkum P G, Franx M, Kriek M, et al. Confirmation of the remarkable compactness of massive quiescent galaxies at $z \sim 2.3$: early-type galaxies did not form in a simple monolithic collapse [J]. *The Astrophysical Journal*, 2008, 677(1): L5.
- Vika M, Bamford S P, Häußler B, et al. Megamorph—multiwavelength measurement of galaxy structure. sérsic profile fits to galaxies near and far [J]. *Monthly Notices of the Royal Astronomical Society*, 2013, 435(1): 623-649.
- Vikram V, Wadadekar Y, Kembhavi A K, et al. Pymorph: automated galaxy structural parameter estimation using python [J]. *Monthly Notices of the Royal Astronomical Society*, 2010, 409(4): 1379-1392.
- Wainscoat R, Freeman K, Hyland A. The optical and near-infrared distribution of light in the edge-on galaxy ic 2531 [J]. *Astrophysical Journal, Part 1 (ISSN 0004-637X)*, vol. 337, Feb. 1, 1989, p. 163-190., 1989, 337: 163-190.
- Walcher J, Groves B, Budavári T, et al. Fitting the integrated spectral energy distributions of galaxies [J]. *Astrophysics and Space Science*, 2011, 331: 1-51.
- Wang C Y, Bochkovskiy A, Liao H Y M. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2023: 7464-7475.
- Weinzirl T, Jogee S, Khochfar S, et al. Bulge n and b/t in high-mass galaxies: constraints on the origin of bulges in hierarchical models [J]. *The Astrophysical Journal*, 2009, 696(1): 411.
- Willett K W, Lintott C J, Bamford S P, et al. Galaxy zoo 2: detailed morphological classifications for 304 122 galaxies from the sloan digital sky survey [J]. *Monthly Notices of the Royal Astronomical Society*, 2013, 435(4): 2835-2860.
- Williams R J, Quadri R F, Franx M, et al. The evolving relations between size, mass, surface density, and star formation in 3×10^4 galaxies since $z = 2$ [J]. *The Astrophysical Journal*, 2010, 713(2): 738.
- Wright S. Correlation and causation [J]. *Journal of agricultural research*, 1921, 20(7): 557.
- Xu Q, Shen S, de Souza R S, et al. From images to features: unbiased morphology classification via variational auto-encoders and domain adaptation [J]. *Monthly Notices of the Royal Astronomical Society*, 2023, 526(4): 6391-6400.
- Yu S Y, Ho L C, Wang J. Spiral structure boosts star formation in disk galaxies [J]. *The Astrophysical Journal*, 2021, 917(2): 88.
- Zamojski M, Schiminovich D, Rich R, et al. Deep galex imaging of the cosmos hst field: A first look at the morphology of $z \sim 0.7$ star-forming galaxies [J]. *The Astrophysical Journal Supplement Series*, 2007, 172(1): 468.

- Zhang Y, Zhao Y. Astronomy in the big data era [J]. *Data Science Journal*, 2015, 14: 11-11.
- Zhu L, van den Bosch R, van de Ven G, et al. Orbital decomposition of califa spiral galaxies [J]. *Monthly Notices of the Royal Astronomical Society*, 2018, 473(3): 3000-3018.

致 谢

“为什么你要去爬珠穆朗玛峰?”

“因为山就在那里。”

三年说长不长，说短不短。一晃眼已是 2024 年春末，居然是我要毕业的日子了。顶着不问世事、稚嫩天真的评价，我在这看似百无一用的学科竟已晃荡了七年的时光。儿时的梦想就像山间的云岚。我于山下仰望时有万千向往，不惧前路的未知，幻想攀至顶峰，脚踏轻雾的随性。而当决意攀爬才发现，快乐的时间太少，痛苦的时间很长。且不说天资有限，处处的困境与解决困境的漫长煎熬，单是攀至云岚所在之处，本幻想着是多少的欢喜，却是身在云岚不可观云岚。只手可摘的愉悦像是镜花水月，映入眼帘的只有更高处的风景。硕士毕业只是漫长长路的一个途经点，凡心所向，素履所往。

牢骚与决心如上所示，下面进入正文环节：

首先感谢我硕士阶段的二位导师：沈世银老师和 Rafael S. de Souza 老师。都说一个师门走不出两样的学生，我对科研的态度，问题的看法受沈老师影响颇深，已然是有沈老师的形状了。Rafael 于我亦师亦友，给予了我非常详细的指导与意见（如周末绝不加班），我们也经常一起探讨与科研无关的有趣话题。虽然二位导师有着迥然不同的风格，但这也帮助了我在冲突中建立自我意识。感谢幸福星团的老师们，你们就像一个大家庭中的家长，给了我许多温暖与关怀。

感谢 ML 组的徐权峰、叶人豪。同样的学术追求让我们走到一起，半夜被 HAL9000 所加热的机房温度就是我们友谊的温度。感谢其他 KKC 的金刚们，包括一起猩猩的钟文心、以及卢家风、穆子豪、邱佳杰、曾琪、赵倩文。感谢广义十五楼大办公室（现 209）的同学们，感谢 DDL 战友项光漫的巴斯克蛋糕。

感谢 Ana 和 Marco，在巴西短暂的三个月以及之后漫长的申博环节给予我的关心与帮助。

感谢研一时在雁栖湖的排球队队友们：陈康安、安邦硕、李子昂、陈中玄、谢世元、张鑫、谭格非、蔡沛沅，顾弘睿，以及胡侯伦。在雁栖湖的一年是我硕士阶段最单纯、最快乐、最无忧无虑的时光。在遇到挫折与艰辛时，和你们在一起的记忆时常在治愈着我。雁栖湖的风永远承载着我们的友谊。

感谢我同年同月同日生的好朋友陈弓羽，我们是老乡，是朋友，是学术上的同道者，是知音。

最后，感谢我的家人们，包括我的父母陈晓城先生、胡敏晓女士；父、母系的亲人们，以及我的对象刘俊杰。没有你们的支持，我也无从追寻学术的自由。

作者简历及攻读学位期间发表的学术论文与研究成果

作者简历：

陈宓，女，1999年6月出生于浙江省温州。

教育简历：

2005年9月至2010年6月，就读于浙江省温州市实验小学；

2010年9月至2014年6月，就读于浙江省温州市第二外国语学校；

2014年9月至2017年6月，就读于浙江省温州市温州中学；

2017年9月至2021年6月，于山东大学空间科学与物理学院，空间科学菁英班，获得空间科学与技术学士学位；

2021年9月至2024年6月，于中国科学院上海天文台，天体物理专业，攻读硕士学位。

已发表（或正式接受）的学术论文：

1. **Chen, M.**, and de Souza, R. S., Xu, Q., Shen, S., Chies-Santos, A. L., Ye, R., ... Cong, Y.(2024). *GalMOSS: GPU-accelerated Galaxy Profile Fitting*. *Astronomy and Computing*, 100825.
2. Xu, Q., and Shen, S., de Souza, R. S., **Chen, M.**, Ye, R., ... Durgesh, R. (2023). *From Images to Features: Unbiased Morphology Classification via Variational Auto-Encoders and Domain Adaptation*. *MNRAS*, 526(4), 6391-6400.

获奖情况：

2024 学年中银科苑优秀学子

2024 学年上海天文台首届研究生学术论坛一等奖

2023 至 2024 学年学业一等奖学金

2022 至 2023 学年三好学生

